B1

(19) Europäisch s Patentamt
Europ an Patent Office
Office uropéen des brevets

(11) EP 1 172 374 A1

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
16.01.2002 Bulletin 2002/03

(51) Int Cl.[7]: **C07K 14/195**, G06F 17/50

(21) Application number: **01306060.3**

(22) Date of filing: **13.07.2001**

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE TR
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **14.07.2000 GB 0017376**
19.09.2000 GB 0229435

(71) Applicants:
- **Medical Research Council**
 London W1B 1AL (GB)
- **The University of Utah Research Foundation**
 Salt Lake City, Utah 84108 (US)

(72) Inventors:
- **Ramakrishnan, Venkatraman, MRC Laboratory**
 Cambridge, Cambridgeshire CB2 2QH (GB)
- **Brodersen, Ditlev Egeskov, MRC Laboratory**
 Cambridge, Cambridgeshire CB2 2QH (GB)
- **Carter, Andrew Philip, MRC Laboratory**
 Cambridge, Cambridgeshire CB2 2QH (GB)
- **Wimberly, Brian Thomas**
 Kalamazoo, MI 49008 (US)
- **Clemons Jr., William Melvon**
 Boston, MA 02115-6091 (US)

(74) Representative: **Brasnett, Adrian Hugh**
Mewburn Ellis York House 23 Kingsway
London WC2B 6HP (GB)

(54) Crystal structure (3 A resolution) of the 30S ribosome and its use

(57) The invention provides an X-ray crystal structure of the 30S ribosome, obtained from *Thermus thermophilus* 30S subunit, having a tetragonal space group $P4_12_12$ with unit cell dimensions of a = 401.4 ± 4.0Å, b = 401.4 ± 4.0Å, c = 175.9 ± 5.0Å. An advantageous feature of the structure is that it diffracts beyond 3Å resolution. The invention also provides a crystal of 30S having the three dimensional atomic coordinates of the 30S ribosome, the coordinates being provided in Tables 1A and 1B. The data may be used for the rational design and modelling of inhibitors for the 30S ribosome, which have potential use as antibiotics.

EP 1 172 374 A1

Description

Field of the Invention

[0001] The present invention relates to the provision of a high resolution crystal structure of the prokaryotic 30S ribosomal subunit, and the use of this structure in drug discovery.

Background of the Invention

[0002] The wealth of information made available through efforts in structural genomics and advances in computation has allowed structure-based drug design to emerge as a valuable tool in medicinal chemistry. In the past combinatorial chemistry, coupled with high-throughput approaches, shifted attention away from the more structure-based methods. Large-scale determination of protein structures is reversing the drug discovery process by starting with the protein structure and using it to identify and design new ligands. It is the integration of structure-based methods, virtual screening, and combinatorial chemistry that will provide the basis for more efficient drug design in the future, significantly reducing the time of the design cycle and the cost per marketed drug. Significant advances have already been made in AIDs, arthritis and cancer and in the treatment of hypertension e.g. captopril.

[0003] Translation of the genetic code occurs on the ribosome, a large nucleoprotein complex that consists of two subunits. In bacteria, the two subunits are denoted 30S and 50S. The 50S subunit contains the catalytic site of peptidyl transferase activity, while the 30S subunit plays a crucial role in decoding messenger RNA. Protein synthesis is a complex, multistep process that requires several extrinsic GTP-hydrolysing protein factors during each of the main stages of initiation, elongation and termination. Despite several decades of work, the molecular details of the process are poorly understood, and the elucidation of the mechanism of translation is one of the fundamental problems in molecular biology today. A recent collection of articles summarizes the state of understanding of the field [1].

[0004] A contribution to this problem was made by Yonath and coworkers, who after nearly a decade of work showed that structures as large as the 50S ribosomal subunit would form crystals that diffract beyond 3Å resolution [2]. Originally, it was not clear that phase information from such a large asymmetric unit could be obtained to high resolution, but the development of bright, tuneable synchrotron radiation sources, large and accurate area detectors, vastly improved crystallographic computing, and the advent of cryo-crystallography have all contributed to making structural studies of the ribosome more tractable. In our work, the use of anomalous scattering from the LIII edges of lanthanides and osmium has also played a critical role in obtaining phases.

[0005] The 30S ribosomal subunit (hereafter referred to as 30S) from *Thermus thermophilus* was originally crystallized by Trakhanov *et al.* in 2-methyl-2,4-pentanediol (MPD) [3] and soon afterwards by Yonath and coworkers in a mixture of ethylbutanol and ethanol [4]. Subsequent work by both groups showed that the MPD crystal form diffracted to about 9-12Å resolution [5, 6]. The diffraction limit of these crystals did not improve beyond 7Å resolution for almost a decade, but more recently both Yonath and coworkers [7, 8] and we [9] obtained crystals of the MPD form that exhibit significantly improved diffraction. However, unlike the crystals obtained by the Yonath group [6], our crystals do not require soaking in tungsten clusters or heat treatment in order to obtain high resolution diffraction.

[0006] We have previously described the structure of the 30S at 5.5Å resolution [9]. We were able to place all seven proteins whose structures were known at the time, infer the structure of protein S20 to be a three-helix bundle, trace the fold of an entire domain of 16S RNA, and identify a long RNA helix at the interface that contains the decoding site of the 30S. Proteins S5 and S7 were also placed in electron density maps of the 30S obtained by Yonath and coworkers.

[0007] The 30S ribosomal subunit is a major target for antibiotics. The ribosome is a useful target for antibiotics since the structure of the 30S is widely conserved between prokaryotes, allowing for broad spectrum antibiotics. However, resistance to current antibiotics is currently a major problem in the field of medicine. There are presently very few new antibiotics available which can be used to treat the highly resistant strains of bacteria such as MRSA (methicilin resistant *Staphylococcus aureus*) which are becoming increasingly widespread.

[0008] Understanding the interaction of antibiotics with the ribosome at the molecular level is important for two reasons. Firstly, antibiotics act by interfering with various aspects of ribosome function. Thus understanding their interaction will help shed light on mechanisms involved in translation. Secondly, a detailed knowledge of antibiotic interactions with the ribosome could aid the development of new drugs against increasingly resistant strains of bacteria. Although antibiotics were characterized several decades ago, a detailed knowledge of their mechanism will in general require a three-dimensional structure of their complex with the ribosome.

[0009] The low (greater than 3Å resolution) crystal structures described above do not provide sufficiently detailed resolution for useful modelling of the crystal structure of the 30S and there is thus a need for a high resolution structure which can be usedusefully in the development of novel therapeutics.

Summary of the Invention

[0010] We have now solved and refined the structure of the 30S at 3Å resolution. The structure contains all of the ordered regions of 16S RNA and 20 associated proteins, and contains over 99% of the RNA sequence and 95% of the protein sequences, with the missing parts being exclusively at the termini of RNA or polypeptide chains. Here we describe the overall architecture and the main structural features of the 30S subunit.

[0011] The refined atomic resolution model of the 30S presented here allows the interpretation of a vast amount of biochemical data on its function in precise structural terms. The structure will also serve as a basis for the interpretation in molecular terms of lower resolution models of various functional states by electron-microscopy or x-ray crystallography. The 30S structure will help produce testable models for various aspects of ribosome function.

[0012] In a first aspect, the present invention provides a crystal of the *Thermus thermophilus* 30S subunit having a tetragonal space group $P4_12_12$ with unit cell dimensions of a = 401.375 Å, b = 401.375 Å, c = 175.887 Å, or more generally a = 401.4 Å, b = 401.4 Å, c = 175.9 Å, but more preferably a = 401.4 ± 4.0Å, b = 401.4 ± 4.0Å, c = 175.9 ± 5.0Å. An advantageous feature of the structure is that it diffracts beyond 3Å resolution. Another feature of the structure is that it was obtained in a method which did not involve soaking crystals in heavy atom (e.g. tungsten or tantalum) clusters or heat activation. Furthermore, it is specifically of the 885-888/910-912 base pairing confirmation of 16S RNA. These features, both singly and in combination all contribute to features of the invention which are advantageous.

[0013] In a second aspect, the invention also provides a crystal of 30S having the three dimensional atomic coordinates of the 30S ribosome. Table 1A provides a set of atomic coordinates of the 30S ribosome. Table 1B provides a set based upon the coordinates of Table 1A but which have been refined further from our data. Reference herein to "Table 1" is a reference to either of Table 1A or 1B (or where the context permits, both). Thus, for example, where it is stated that the invention refers to computer readable media with "atomic coordinate data according to Table 1 recorded thereon", this means that the media has either the data of Table 1A, or the data of Table 1B, or both, recorded thereon.

[0014] We have also observed that 30S crystals do not contain the S1 subunit protein. In our studies, we have found that by selectively removing this protein prior to crystallization, we have been able to obtain the improved resolution described herein. Although the atomic co-ordinates provided in Table 1 below allows those of skill in the art to bypass the need to undertake the crystallization of the 30S, this crystallization method nonetheless forms a further aspect of the invention.

[0015] Accordingly, there is provided a method for crystallizing a the 30S subunit to obtain a high resolution structure of a 30S subunit, which method comprises providing a 30S subunit, selectively removing the 31 subunit therefrom and crystallizing the 30S.

[0016] In a further aspect, the present invention provides a method for identifying a potential inhibitor of the 30S comprising the steps of:

a. employing a three-dimensional structure of 30S, or at least one sub-domain thereof, to characterise at least one active site, the three-dimensional structure being defined by atomic coordinate data according to Table 1; and
b. identifying the potential inhibitor by designing or selecting a compound for interaction with the active site.

[0017] In a further aspect, the present invention provides computer readable media with either (a) atomic coordinate data according to Table 1 recorded thereon, said data defining the three-dimensional structure of 30S or at least one sub-domain thereof, or (b) structure factor data for 30S recorded thereon, the structure factor data being derivable from the atomic coordinate data of Table 1.

Description of the Drawings.

[0018]

Figure 1 shows the secondary structure of the 30S ribosome.

Figure 2 is Table 1A and 1B.

Detailed Description of the Invention.

Definitions.

[0019] The term "sub-domain" includes the following:

(a) and elements selected from the following:

at least one complet lement of s condary structure, i. . an alpha helix or a b ta sheet, or RNA helix, as describ d in the detailed description b low;
a group of two or more such l m nts which interact with each oth r;
at least on subunit protein;
a subgroup of subunit proteins, for xampl a group which includes two or mor prot ins which are found to interact with each other;
any of the above, when being protein or element thereof being used in conjunction with all or part of the 16S RNA structure associated with said elements or proteins;

(b) a space of volume defining a region around any one particular atom of interest (e.g. an atom involved in binding to an antibiotic), the volume being less than the total volume of the tetragonal space of the complete crystal. For example, the coordinates of atoms in a volume of from about 500 to about 15,000Å^3 may be selected and used for the present invention. Such a space may be a sphere having a diameter of from about 10Å to about 30Å, centred around a point of interest; and

(c) a collection of at least 10, e.g. at least 25 such as at least 50, more preferably at least 100, even more preferably at least 500 atoms and most preferably at least 1000 atoms defined by the coordinates of Table 1, wherein at least 2 of said atoms, and preferably at least 50% of said atoms of the collection are located within 50Å of each other.

[0020] An "active site" of the 30S is any part of this structure involved in tRNA or mRNA binding, factor binding or translocation. This includes regions responsible for binding initiation factors, elongation factor G or release factors, regions which are target sites for regulation by co-factors, phosphorylation or acetylation, and regions responsible for interaction with the 50S ribosome. In also includes regions which change conformation during translocation or protein synthesis, particularly one or more of the 16S RNA helixes 18, 27, 34 and 44.

[0021] Particular regions of the 30S include antibiotic binding regions. Other regions include the three tRNA binding sites, i.e. the aminoacyl (A), peptidyl (P) and (exit) E sites.
Other active sites are those which undergo movement during translocation of tRNAs from the A to P sites and the P to E sites. Regions further include any one of the subunit proteins S2 to S20 and THX, including any of the individually identified subunit proteins in the accompanying examples.

[0022] By "fitting", is meant determining by automatic, or semiautomatic means, interactions between one or more atoms of an potential inhibitor molecule and one or more atoms or binding sites of the 30S, and calculating the extent to which such interactions are stable. Various computer-based methods for fitting are described further herein.

[0023] By "root mean square deviation" we mean the square root of the arithmetic mean of the squares of the deviations from the mean.

[0024] "Computer readable media" refers to any media which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media such as floppy discs, hard disc storage medium and magnetic tape; optical storage media such as optical discs or CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media.

[0025] A "computer system" refers to the hardware means, software means and data storage means used to analyse the atomic coordinate data of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means and data storage means. Desirably a monitor is provided to visualise structure data. The data storage means may be RAM or means for accessing computer readable media of the invention. Examples of such systems are microcomputer workstations available from Silicon Graphics Incorporated and Sun Microsystems running Unix based, Windows NT or IBM OS/2 operating systems.

Table 1.

[0026] The coordinates of Table 1 provide a measure of atomic location in Angstroms, to a third decimal place. In order to use the information in these Tables for the purposes described herein as being aspects of the present invention, these coordinates may be varied by up to ± 1.0, such as by up to ± 0.7, preferably no more than up to ±0.5 Angstroms, without departing from the scope of the invention.

[0027] Furthermore, varying the relative atomic positions of the atoms of the structure so that the root mean square deviation of the 16S RNA or S2-S20 protein backbone atoms is less than 1.5Å (preferably less than 1.0Å and more preferably less than 0.5Å) when superimposed on the coordinat s provided in Tabl 1 for thes structur s, will g n rally result in a structure which is substantially the sam as th structure of Tabl 1 in t rms of both its structural characteristics and potency for structure-bas d drug design of 30S ligands.

[0028] Thus for th purposes described herein as being aspects of the present inv ntion, it is within th scop of th

invention if: the Table 1 coordinates are transposed to a different origin and/or axes; the relative atomic positions of the atoms of the structure are varied so that the root mean square deviation of conserved residue backbone atoms is less than 1.5Å (preferably less than 1.0Å and more preferably less than 0.5Å) when superimposed on the coordinates provided in Table 1 for the conserved residue backbone atoms; and/or the number and/or positions of water molecules is varied. Reference herein to the use of the coordinates of Table 1 thus includes the use of coordinates in which one or more individual values of the Table are varied in this way.

[0029] Table 1 includes coordinates of two zinc ions, together with 202 other ions which are not identified, although, while not wishing to be bound by any one theory, are believed to be selected from cobalt and magnesium. Some or all of these ions may optionally be discarded from Table 1 when using the data. The table also lists the coordinates of a 26 amino acid peptide, Thx, as well as a 6 nucleotide fragment of mRNA, NNNUCU, designated as molecule X. Both the coordinates of both these molecules may likewise optionally be discarded, i.e. so that the coordinates of the 16S RNA and the proteins S2 to S20 alone are modelled and used in applications of the invention.

[0030] There are a few N- or C-terminal sequences of the S2 to S20 proteins which were not resolved in the structure of Table 1, together with a some of the 5' and 3' residues of the 16S RNA. These are not essential for the purposes of the present invention, but are listed in Table 2 for completeness. Those of skill in the art may, if desired, wish to adapt the structures provided by the coordinate of Table 1 by modelling in one or more of the amino acids or nucleotides of Table 2.

[0031] This methodology provides those of skill in the art a means to provide 30S crystals of *T.thermophilus.* The conservation of ribosome structure, particularly regions of structure essential for function, between prokaryotes, for example prokaryotes which are human pathogens, such as Staphylococcus spp, and the like, allows the structure herein to be useful in the provision of anti-bacterial agents in general. Thus, the structure may be used to solve 30S subunits by the technique of molecular replacement. In such a method, x-ray diffraction data are obtained from crystals of a 30S subunit from another species, e.g. a species of a bacteria pathogenic to humans. The coordinates of Table 1 may be used to find the orientation of the unknown molecule in the crystal, and electron density maps calculated. These maps can then be interpreted with the sequence of the species in question, and the coordinates of our 30S structure can be used to help and speed interpretation. In this way, the structure of our 30S facilitates the determination of structures of 30S subunits and whole ribosomes from other organisms.

[0032] Accordingly, the invention provides a method for the determination of the structure of a bacterial 30S from a species other than *T. thermophilus* which method comprises:

(a) crystallising the 30S of said species to obtain a crystal;
(b) performing X-ray crystallography on said crystal to obtain X-ray diffraction data;
(c) providing the structure data of Table 1; and
(d) using molecular replacement to calculate an electron density map of the 30S.

[0033] In such a method the 30S may be prepared by removal of the S1 subunit, as described herein.

[0034] The electron density map obtained may then be used to calculate the atomic coordinate date of the 30S. The atomic coordinate date thus obtained may be used to for the design and analysis of new and specific ligands for 30S as described herein.

The 30S crystal structure.

[0035] The high resolution structure provided herein provides a crystal with unit cell dimensions which are provided in the accompanying table to 3 decimal places, i.e. a=b= 401.375, c = 175.887 Å. However, those of skill in the art wishing to reproduce the crystallization described herein and obtain such crystals will appreciate that a degree of experimental variability and error will mean that crystals of the invention will be obtained with a unit cell dimension within, but not exactly corresponding to, this size. Thus crystals of the invention may generally be defined as having unit cell dimensions of a = 401.4 ± 4.0Å, b = 401.4 ± 4.0Å, c = 175.9 ± 5.0Å, preferably a = 401.4 ± 1.0Å, b = 401.4 ± 1.0Å, c = 175.9 ± 2.0Å, preferably a = 401.4 ± 0.7Å, b = 401.4 ± 0.7Å, c = 175.9 ± 1.4Å, and more preferably a = 401.4 ± 0.2Å, b = 401.4 ± 0.2Å, c = 175.9 ± 0.4Å. These unit cell sizes are believed to define a novel and more highly resolved unit cell size than has previously been possible in the art.

Production of crystals.

[0036] To obtain crystals according to the present invention, we have found that selective removal of the S1 subunit protein is advantageous. A suitable method for the selective removal of the S1 subunit protein is by the use of a hydrophobic interaction chromatography column (poros-ET). 30S ribosomal subunits lacking the S1 subunit may suitably be separated from those containing the S1 subunit by running a column using a reverse ammonium sulfate gradient

from 1.5M to 0.5M, with 20mM Hepes, pH 7.5, and 10mM acetat . The 30S subunits lacking S1 are elut d first, giving th first major peak. During elution of th 30S peak th ammonium sulfate concentration is maintain d at a constant lev l. Onc th 30S p ak has eluted th ammonium sulfate concentration is then furth r r duc d to elut th 30S + S1 fraction.

[0037] An alternative method for th s lective removal of th S1 subunit protein is by pr parativ gel lectrophoresis. Gel electrophoresis may suitably be carried out by first preparing and mixing a 3% acrylamide, 0.5% agaros cylindrical gel, and pouring this gel into a BioRad Prep Cell. 30S ribosomal subunits are then loaded onto the gel and continuously eluted as they emerge form the other end of the gel. The 30S fraction lacking the S1 subunit comes off first, giving the first major peak. The 30S + S1 fraction gives the trailing peak (or shoulder) and can be discarded.

[0038] Once the S1 is removed, the crystals may be formed, using suitable conditions. These include the use of 13-17% v/v methyl-2,4-pentanediol in the presence of 200-300 (e.g. about 250) mM KCl, 50-100 (e.g. about 75) mM ammonium chloride, 15-30 (e.g. about 15 or about 25) mM $MgCl_2$ at a pH of 6.0 - 7.5 (e.g about pH 6.3 - 6.7 such as pH 6.5) in 50 - 150 (e.g. about 100) mM sodium or potassium cacodylate or MES (2-(N-morpholino)ethane sulphonic acid).

[0039] In a particular aspect, the conditions may comprise the use of 250 mM KCl, 75 mM NH_4Cl, 25 mM $MgCl_2$, 6 mM 2-mercaptoethanol in 0.1 M potassium cacodylate or 0.1 M MES (2-N-morpholinoethanesulfonic acid) at pH 6.5 with 13-17% MPD as the precipitant.

[0040] The crystals may be grown by any suitable method known as such to those of skill in the art. Suitably, the crystals may be grown over a period of 4-8 weeks at about 4°C. The structure of the crystals so obtained may be resolved, and crystals which resolve to a resolution of at least 3Å selected. Crystals which resolve to a resolution of at least 3Å obtainable by such a method are a further aspect of the invention.

Use of structure of Table 1.

[0041] The determination of the three-dimensional structure of 30S provides a basis for the design of new and specific ligands for 30S. For example, knowing the three-dimensional structure of 30S, computer modelling programs may be used to design different molecules expected to interact with possible or confirmed active sites, such as binding sites or other structural or functional features of 30S.

[0042] The high resolution model of the 30S provided by Table 1 may be used to examine and determine the binding of antibiotics known to target this ribosome subunit. Such antibiotics include paromomycin, streptomycin, spectinomycin, tetracycline, pactamycin and hygromycin B.

[0043] A candidate ligand, particular one which acts as an inhibitor molecule may be any available compound. A number of commercial sources of libraries of compound structures are available, for example the Cambridge Structural Database.

Such libraries may be used to allow computer-based high throughput screening of many compounds in order to identify those with potential to interact with the active site of a ribosome.

[0044] More specifically, a potential ligand capable of modulating 30S activity can be examined through the use of computer modelling using a docking program such as GRAM, DOCK, or AUTODOCK (see Walters et al., *Drug Discovery Today,* Vol.3, No.4, (1998), 160-178, and Dunbrack et al., *Folding and Design,* 2, (1997), 27-42) to identify potential ligands of 30S. This procedure can include computer fitting of potential ligands to 30S or a subdomain thereof to ascertain how well the shape and the chemical structure of the potential ligand will bind to the enzyme.

[0045] Also computer-assisted, manual examination of the active site structure of 30S may be performed. The use of programs such as GRID (Goodford, *J. Med. Chem.*, 28, (1985), 849-857) - a program that determines probable interaction sites between molecules with various functional groups and the enzyme surface - may also be used to analyse the active site to predict partial structures of ligands for the site.

[0046] Computer programs can be employed to estimate the attraction, repulsion, and steric hindrance of the two binding partners (e.g. the 30S and a potential ligand). Generally the tighter the fit, the fewer the steric hindrances, and the greater the attractive forces, the more potent the potential ligand since these properties are consistent with a tighter binding constant. Furthermore, the more specificity in the design of a potential ligand, the more likely it is that the ligand will not interact with other proteins as well. This will tend to minimise potential side-effects due to unwanted interactions with other proteins.

[0047] Having designed or selected possible binding ligands, these can then be screened for activity. Consequently, the method preferably further comprises the further steps of:

obtaining or synthesising the pot ntial ligand; and
contacting the potential ligand with 30S to det rmin the ability of the potential ligand to interact with 30S.

[0048] More pref rably, in latter step th potential ligand is contacted with 30S under conditions to determin its

function, for example in a cell free translation system.

[0049] Instead of, or in addition to, performing such an assay, the method may comprise the further steps of:

obtaining or synthesising said potential ligand;

forming a complex of 30S and said potential ligand; and

analysing said complex by X-ray crystallography to determine the ability of said potential ligand to interact with 30S. Detailed structural information can then be obtained about the binding of the potential ligand to 30S, and in the light of this information adjustments can be made to the structure or functionality of the potential ligand, e.g. to improve binding to the active site. Steps c. to e. may be repeated and re-repeated as necessary.

[0050] Another aspect of the invention includes a compound which is identified as an ligand of 30S by the method of the above aspects of the invention.

[0051] The present high resolution structure of 30S provides a means to determine the location of binding of antibiotics, as well as the interactions at the location between 30S and the antibiotic. Such antibiotics include paromomycin, streptomycin spectinomycin, tetracycline, pactamycin and hygromycin B. The high structure of Table 1 may be used to model the binding to 30S of these, other antibiotics and other ligands. Thus in another aspect, the invention provides a method of analysing a 30S-ligand (wherein "ligand" includes an antibiotic) complex comprising the steps of (i) cocrystallising the 30S with the ligand or soaking the ligand into crystals of the 30S; (ii) collecting X-ray crystallographic diffraction data from the crystals of the 303-ligand complex and (iii) using the three-dimensional structure of 30S of Table 1, or at least one sub-domain thereof, to generate a difference Fourier electron density map of the 30S-ligand; and (iv) modelling the ligand in the difference Fourier electron density.

[0052] Therefore, 30S-ligand complexes can be crystallised and analysed using X-ray diffraction methods, e.g. according to the approach described by Greer et al., *J. of Medicinal Chemistry,* Vol. 37, (1994), 1035-1054, and difference Fourier electron density maps can be calculated based on X-ray diffraction patterns of soaked or co-crystallised 30S and the solved structure of uncomplexed 30S. These maps can then be used to determine the structure of the ligand bound to the 30S and/or changes the conformation of 30S.

[0053] Data obtained from a ligand bound to 30S may be used to improve the ligand, for example by adding or removing functional groups, substituting groups or altering its shape to obtain improved candidates, which may then be screened, solved in complex as described herein above, in an iterative process.

[0054] Electron density maps can be calculated using programs such as those from the CCP4 computing package (Collaborative Computational Project 4. The CCP4 Suite: Programs for Protein Crystallography, *Acta Crystallographica,* D50, (1994), 760-763.). For map visualisation and model building programs such as "O" (Jones et al., *Acta Crystallograhy,* A47, (1991), 110-119) can be used.

[0055] By providing such computer readable media, the atomic coordinate data can be routinely accessed to model 30S or a sub-domain thereof. For example, RASMOL is a publicly available computer software package which allows access and analysis of atomic coordinate data for structure determination and/or rational drug design.

[0056] On the other hand, structure factor data, which are derivable from atomic coordinate data (see e.g. Blundell et al., in *Protein Crystallography,* Academic Press, New York, London and San Francisco, (1976)), are particularly useful for calculating e.g. difference Fourier electron density maps.

[0057] In another aspect, the present invention provides systems, particularly a computer systems, intended to generate structures and/or perform rational drug design for 30S or 30S ligand complexes, the systems containing either (a) atomic coordinate data according to Table 1, said data defining the three-dimensional structure of 30S or at least one sub-domain thereof, or (b) structure factor data for 30S, said structure factor data being derivable from the atomic coordinate data of Table 1.

[0058] Mutant strains resistant to the action of these antibiotics can arise through mutation of a protein subunit of the 30S or through mutation or modification in the 16S RNA (e.g. 2'O-methylation), or modification (e.g. acetylation) of the antibiotic). The sites of mutations in some cases are known or can be identified. Where such sites are identified through, for example, primary sequence data, the invention provides a means to model the structure of the mutants.

[0059] There is thus provided a method which comprises providing the structure of the 30S ribosome of Table 1, changing one amino acid or nucleotide of said structure to provide a mutant 30S, and modelling the structure of the mutant 30S to provide a structure of the mutant. The mutant may be used in the manner described above for the wild type, e.g. stored in computer readable form, modelled to provide ligands, and the like. The modelling may be based upon the predicted behaviour of the atoms of the changed amino acid based upon its interaction with the surrounding atoms in the model provided herein.

[0060] This process may be iterative, e.g. to produce successive mutations into the 30S structure, for example 2, 3, 4, or 5 to 10 mutations.

[0061] Regions of 30S which may be subject to this aspect of the invention include those regions identified in the accompanying examples as regions of the 30S involved in ribosome function.

[0062] In a further aspect, the present invention provides a means to solve or interpret electron density maps of the whole 70S ribosome at low or high resolution, and thus solve the structure of the whole 70S ribosome.

[0063] In particular, the invention provides a method for the determination of the structure of a bacterial 70S ribosome which method comprises

(a) crystallising the 70S of said species to obtain a crystal;
(b) performing X-ray crystallography on said crystal to obtain X-ray diffraction data;
(c) providing the structure data of Table 1; and
(d) using molecular replacement to calculate an electron density map of the 70S.

[0064] The invention is illustrated below by the following examples, their accompanying Figures and Tables. In Table 1 there is shown in each row Atom number, element type, residue (amino acid, nucleotide, etc), number in molecule (for proteins N to C terminal direction, for nucleic acid 5' to 3' direction), X, Y and Z co-ordinates, occupancy, B factor ($Å^2$) and an identifier for the member of the 30S (e.g. for the subunits in the format "ASn" where A is an arbitary letter, different for each member, S is the subunit and n is the subunit number; and for the 16S as "A16S").

[0065] Throughout the accompanying example, we use the numbering system for *E. coli* 16S RNA, as well as the standard helix numbering, denoted H1-H45, for the secondary structure elements [19] with some modifications as shown in Figure 1. The most significant differences between the *E. coli* and *T. thermophilus* sequences are a shorter H6 and H10, and insertions in H9 and H33a. Any insertions in *T. thermophilus* relative to E. coli are indicated in the coordinates with an insertion letter after the nucleotide number, following the practice for tRNA.

MATERIALS AND METHODS

Crystallization of the 30S.

[0066] Because we observed that the 30S crystals completely lacked ribosomal protein S1, care was taken to remove S1 selectively from the 30S prior to crystallization. Crystals were obtained in 13-17% MPD over a range of pH in the salt and magnesium conditions described by Trakhanov et al [3]. The crystals were largest and most reproducibly obtained at a pH of 6.5 in 0.1 M cacodylate or MES buffer. Crystals took approximately 6 weeks at 4°C to grow to their maximum size. The largest crystals, which were required for high resolution data collection, grew to a size of 80-100 □80-100 □200-300 microns. The activity of redissolved crystals in poly(U)-directed protein synthesis was comparable to that of freshly isolated 30S subunits.

Data collection.

[0067] Crystals were transferred to 26% MPD by vapor diffusion in two steps over a period of 6 days. All crystals (except for those soaked in osmium hexammine or osmium pentammine) also contained 1 mM cobalt hexammine in the cryoprotectant. Crystals were flash-cooled by plunging into liquid nitrogen, and data collection was done in a cryostream at 90-100 K.

[0068] A large fraction of crystals was screened at beamlines 9.6 or 14.1 at the SRS at Daresbury Laboratories, using two short exposures at least 40 degrees apart. These crystals were then analyzed for diffraction limits, cell dimensions and mosaic spread. Only crystals of similar cell dimensions and with reasonable mosaic spread were used for data collection.

[0069] Potential derivatives were screened on beamlines X25 at the NSLS at Brookhaven National Laboratory and BM-14 at the ESRF (Grenoble). Data to about 4.5Å were obtained from X25. High resolution data were collected at SBC ID-19 at the APS in Argonne National Laboratory, and ID14-4 at the ESRF. In all cases, derivative data were collected at the peak of the fluorescence at the LIII edge to maximize anomalous differences. At X25 and SBC ID-19, the kappa goniostat was used to rotate precisely about a mirror plane so that small anomalous differences could be measured accurately. Each crystal typically yielded 3-10 degrees of data. Data were integrated and scaled using HKL-2000 [10].

Structure determination.

[0070] Previously determined phases at 5.5 Å [9] were used to locate heavy atom sites using anomalous difference Fourier maps. Initially, these sites were used for phasing to 3.35 Å using the program SOLVE [11], followed by density modification with SOLOMON [12], using the procedure implemented in SHARP [13]. Optimization of the various parameters in the procedure was required to obtain interpretable maps. The RNA and some of the proteins were built using the SOLVE maps. The sequence of Thermus thermophilus 16S RNA [14] was used for the structure. For proteins,

a combination of previously published sequences and new ones from the Göttingen Thermus genome sequencing project were used. Improved maps were obtained by calculating experimental phases to 3.2Å using SHARP followed by density modification and phase extension to 3.05Å with DM [15]. The improved maps allowed us to build all the ordered parts of the structure. The model was built using O [16], and refined using the program CNS [17]. Maximum likelihood refinement was used, initially with both amplitudes and experimental phase probability distributions to 3.35Å, and subsequently with amplitudes to 3.05Å.

RESULTS

[0071] The 30S subunit from Thermus thermophilus consists of a 1522 nucleotide 16S ribosomal RNA [14] and 21 associated proteins, of which 20 have known counterparts in E. coli. Protein S21 is not present in Thermus, and protein S1 has been removed from the 30S prior to our crystallization. In addition, a 26 residue peptide, Thx, is present in Thermus 30S subunits [18].

[0072] Experimentally phased maps clearly showed main chain density for RNA and protein, individual bases (which were often of sufficient quality to distinguish purines from pyrimidines), and large well-ordered side chains of proteins. These maps were used to build 16S RNA and the previously unknown proteins S2, S3, S9, S10, S11, S12, S13, S14 and Thx. In addition, regions that were disordered in isolated structures or had changed significantly were also built. This often consisted of significant portions of the N- and C-terminal tails of the proteins, sometimes including entire domains that were unfolded in isolation. Proteins with small cores and long loops, such as S16 and S17, had to be substantially rebuilt, since these loops were generally disordered in the solution NMR structures. Finally, the entire structure was rebuilt after an initial round of refinement. Our current model consists of nucleotides 5-1511 of Thermus thermophilus 16S RNA (corresponding to 5-1534 of E. coli 16S RNA) and all of the ordered regions of the associated 20 proteins. The current model has been refined against 3.05 Å data with a conventional R-factor of 0.213, a free R-factor of 0.256 and good geometry. For the proteins, 94% of the residues were in the core or allowed regions of the Ramachandran plot, 3.9% in the generously allowed region and 1.8% in the disallowed region.

16S RNA

[0073] The secondary structure of 16S ribosomal RNA contains forty-five double helices connected by short single-stranded segments. In the crystal structure, many of these helices are coaxially stacked with a helix neighboring in the sequence. There are 13 groups of coaxially stacked helices and 23 unstacked helices in 16S rRNA, for a total of 36 helical elements. There are three different types of helix-helix packing. Most of the helical elements are packed in a minor groove to minor groove fashion, which often requires distortions from canonical A-form helical geometry in one of the two helices. Adenosines from internal loops or from hairpin loops often mediate docking against an A-form double helix, with a dense network of base-2' OH and 2' OH - 2' OH hydrogen bonds stabilising the packing. Less often, helix-helix packing occurs in a different mode, by insertion of a ridge of phosphates into a complementary minor groove of another helix. This packing mode is stabilized by hydrogen bonds between the ridge of phosphate oxygens and a layer of 2' OH and guanine base NH_2 groups. These guanine N2 groups are often made more accessible by the geometry of G-U pairs, which places this moiety farther into the minor groove than do Watson-Crick pairs. Finally, the rare end-on mode of interhelical packing uses a purine base to mediate the perpendicular packing of one helix against the minor groove of another helix. All three modes of helix-helix packing are further stabilized by idiosyncratic interactions between double-helical RNA and short non-helical RNA segments. Small bulges of one to three nucleotides are often found to pack either between helices or in the major groove of a helix.

The 5' domain (fpd).

[0074] The fpd of 16S RNA contains 19 double helices, arranged as 7 groups of coaxially stacked helices and 5 unstacked helices, for a total of 12 double-helical elements packed tightly together. The result is a wedge-shaped mass of RNA that tapers to a single layer of double helices near the top of the domain. Like the other domains, the fpd is rather longer along the subunit interface than in the perpendicular direction.

[0075] The fpd can be divided into three subdomains, roughly corresponding to the upper, lower, and middle thirds of the secondary structure of the fpd. These subdomains make up the top and left-hand, the middle, and the lower right-hand sides of the body, respectively, in the view from 50S. The upper subdomain is a nearly planar arrangement of four helical elements (H16/H17, H4/H15, H1/H3, and H18). The H16/H17 stack forms the left-hand border of the body as viewed from 50S. This stack is almost 120 Å long, with H16 making contact with the head and H17 reaching the bottom of the subunit. Internal loops in both helices contain S-turns, which are used to modulate the position of the phosphate backbone in the case of H17, or to create an extended minor groove surface for helix-helix docking in the case of H16. The H4/H15 stack points towards the bottom of the subunit, with H15 well-packed against H17. The H1/H3

stack is bent by the conserved bulge at position 31, which results in the proximal end bending horizontal and the terminal end pointing up to the head. The fourth helical element is H18, which is sharply bent to accommodate the 530 pseudoknot, defined by the unstacked helices 505-507/524-526 (H18.2) and 521-522/527-528 (H18.1). H18 is wedged between the other two upwards-pointing elements of the upper subdomain, H1/H3 and H16. The 530 pseudoknot packs against the central pseudoknot at the H18.1 - H1 interface.

[0076] The middle subdomain contains four helical elements (H5, H6, H12/H6A, and H13/H14) that form a layer between the upper and lower subdomains in the centre of the body. There are relatively few packing interactions within the subdomain, and several of its helices pack against the upper subdomain on one side and the lower subdomain on the other. Thus at the bottom of the subunit, the conserved root of H6 packs against H8 (lower subdomain) on one side and H15 (upper subdomain) on the other side. Similarly, the H12/H6A stack packs against H4 (upper subdomain) and H7 (lower subdomain). H12/H6A also packs against H5 and the 117 loop, which pack against elements from the upper and lower subdomains, respectively. H5 is well-packed against H15 and the 117 loop stacks with the root of H11. H5 also packs against the H13/H14 stack in the phosphate ridge-minor groove manner. H13/H14 interacts with two different regions of the lower subdomain. The conserved UACG hairpin loop at the end of H14 packs against the 160 GAAA hairpin from H8 while the large conserved hairpin at the end of H13 interacts with H7. This hairpin loop also makes many interactions with elements from the middle subdomain.

[0077] The lower subdomain is a collection of three helical elements that form an open saddle-shaped structure in the lower right-hand corner of the body. The H8/H9 stack stretches from the back of the subunit to the front, with the conserved 160 GAAA hairpin pointing toward the 50S subunit. It packs tightly against the H7/H10 stack at the 4-way junction that joins them, and again at a Thermus-specific interaction between insertions at nucleotides 190 and 129. The H7/H10 stack also makes weak interactions with H15 and H17 from the upper subdomain at the bottom of the subunit. H11 contains two sharp bends that allow its conserved terminal hairpin loop to pack against H7. Both bends are stabilized by short-range minor-groove to minor-groove packing contacts.

The central domain (cd).

[0078] The cd is the RNA component of the platform. Its fold based on our previous 5.5 Å structure [9] is in excellent agreement with our current structure. It contains nine helical elements folded into a W-shape in the 50S view. Two long single-stranded segments of RNA, the 570 and 820 loops, are also important structural elements. The domain is dominated by the long stack of H21/H22/H23, which forms the U-shaped perimeter of the domain. H21 is the only component of the left-hand arm of the W, while H22 and H23 form the base of the right-hand side. The right-hand arm of the W consists of H23B and H24A whose conserved hairpin loops are tightly packed. This arrangement requires sharp bends between H23 and H23B, and between H24 and H24A. The H23/H23B bend is stabilized by short-range minor groove-minor groove packing interactions. The H24/H24A bend is more unusual in that the bend is towards the major groove, which places a ridge of H24A phosphates in the major groove of H24. This major-groove bend is stabilized partly by short-range base-base and base-backbone interactions in the major groove of the bend, and partly by long-range interactions between the bent H24/H24A minor groove and the minor groove of H23. The heart of the central domain is the thicker middle arm of the W, which contains six helical elements (H20, H19/H25, H24, H26/H26A, H27, and H23B) and the 570 and 820 loops. On the left-hand side of the arm, the H26/H26A stack packs tightly against H22, the base of H25, and the 570 loop. The H25/H19 stack packs well with H20 and with the 570 loop. On the right-hand side of the central arm of the W, H23A packs well with H22, the 820 loop stacks on H24, and H24 packs well with the conserved GCAA hairpin loop of H27. In the centre of the arm, H23A packs with H26 in the phosphate ridge-minor groove manner, and the conserved H23A GAAG hairpin loop packs against H20. The 820 loop also interacts with H20, H25, and the 570 loop.

The 3' major domain (tmd).

[0079] The 3' major domain (tpd) is the RNA component of the head of the 30S subunit. From the 50S view, the left-hand side of the head tapers to a beak made of RNA on the 50S side and protein on the solvent side. Like the other domains, the tpd is relatively thin in the direction perpendicular to the intersubunit interface. The tpd consists of fifteen helical elements, most of which do not stack on a neighboring helix, in contrast to the extensive stacking of neigboring helices seen in the fpd and the central domain. The tpd can be divided into three subdomains, which correspond to the upper, middle, and lower portions of the tpd secondary structure. The upper subdomain is an extended structure in the part of the head farthest from the 50S subunit, and makes relatively few packing contacts with RNA from the other head subdomains. The lower and middle subdomains are more globular and are more intimately packed together, and make up the front-right and front-left portions of the head, respectively. The middle subdomain includes the RNA portion of the beak.

[0080] The upper subdomain contains three helical elements that make up a well-separated structure on the solvent

side of th head. Th subdomain is dominated by the H35-H36-H38-H39 stack, which str tches from th top to th bottom of the head. Th other two helical el ments of this subdomain ar H37 and H40, which pack well with each other and loosely with th H35-H36-H38-H39 stack. Th H37-H40 pack is mediated by a s micons rv d GAAA hp in H40 with adjac nt G-C pairs in H37.

[0081] Th smaller middle subdomain is xtended and contains only four h lical el m nts, H32, H33/H33A, H33B and H34. Two of these (H33/H33A and H33B) form the Y-shaped RNA component of the beak. The H33/H33A stack points to the left in the 50S view while H33B points to the right, with its terminal conserved GNRA hairpin loop packed against H32, the covalent connection between the beak and the lower subdomain. H32 in turn packs against the H33-H34 junction as well as the 980 loop in the lower subdomain. With the exception of a small packing interaction with H32, the irregular H34 makes only long-range and somewhat tenuous packing interactions. The first is with H31 in the lower subdomain, an unusually weak minor-groove to minor groove packing. The second interaction is an unusual end-on packing interaction with the minor groove of the H34/H35/H38 junction in the upper subdomain.

[0082] The lower subdomain contains almost half of the tpd RNA and contains seven helical elements (H28/H29, H30, H31/980 loop, H41, H41A, H42 and H43) intimately packed into a globular mass. Helices 42 and 43 are arranged in an approximately parallel fashion at the centre of the fold, and each interacts with at least three of the other helical elements. Helices 42 and 43 dock together by means of a minor-groove to minor-groove packing of their conserved hairpin loops. On the solvent side of the H42/H43 pair, H41 packs with both H42 and H43, while the terminal GCAA hairpin loop of H41A packs against H42. This arrangement requires a sharp bend between H41 and H41A, whose minor grooves pack against each other at the bend. The H43-H41 pack is made more extensive by an underwound A-rich internal loop in H41. On the 50S side of the central H42/H43 pair are H29, H30, H31 and the 980 loop. H43 is well-packed with H29 and makes weaker interactions with H30 and the 980 loop, while H42 is well-packed with H30 and the 980 loop. The H42-H30 pack is mediated by successive conserved G-A pairs at the base of H42. The H43-H29 pack is mediated by a conserved S-turn at the base of H43. An S-turn also mediates the packing of H42 with H41. H31 is a peripheral element of the subdomain, packing well only with H30, but also packs with H34 from the middle subdomain.

The 3' minor domain.

[0083] The 3' minor domain consists of just two helices at the subunit interface. H44 is the longest single helix in the subunit, and stretches from the bottom of the head to the bottom of the body. It projects prominently from the body for interaction with the 50S subunit. H45 is approximately perpendicular to H44, with its conserved GGAA hairpin loop packed against H44 and available for interaction with the large subunit.

PROTEINS IN THE 30S AND THEIR INTERACTION WITH 16S RRNA

[0084] The current structure includes all of the 30S proteins except S1. The proteins generally consist of one or more folded domains, about half of which were known from previous work on isolated proteins. However, nearly all of the proteins contain extended termini or loops which interact intimately with RNA and were disordered in the isolated structures. Although most of the proteins form intimate contacts with ribosomal RNA, there are also protein-protein interactions such as those seen in the S4-S5-S8 and S3-S10-S14 clusters.

Proteins in the central domain (S18, S11, S8, S15).

[0085] **S18:** S18 in the 30S consists of residues 19-88. It consists of two helices, and a third helical element formed by two short turns from different parts of the structure that stack end-to-end. These helices together form a hydrophobic core. The C-terminus interacts with S11.

[0086] **S11:** S11 is a new structure and consists of two helices packed against a sheet, a type of fold seen in many ribosomal proteins. The sheet packs against the minor groove of the 690 loop (H23), and has a C-terminal extension that interacts with the C-terminal extension of S18 and also with the 790 loop (H24). Thus S11 stabilizes folding of the platform, by binding to both H23 and H24 near the tip of the platform.

[0087] **S8:** S8 binds near the H20/H21/H22 three-way junction and makes extensive interactions with H21 and H25. We now have molecular details of these interactions. In particular, two loops from S8 (87-92 and 112-118) wrap around the bulged bases 641-642 which were known to be required for high affinity binding of S8 [20, 21]. The N-terminus of the protein also packs against th minor groov of th 825 st m (H25), thus helping th folding of th central domain. R sidues K55 on S8 and 653 on RNA are next to each other as would be xpected from crosslinking [22]. Th xtension in Thermus S8 of th loop 69-76 packs against S2 from a symm try related molecul .

[0088] **S15:** S15 binds between H20 and H22 n ar the three-way junction.

The 5' domain binding proteins S17, S16 and S20.

[0089] **S17:** Although originally thought to b exclusively a 5' domain binding prot in, S17 also binds n ar th H20/H21/H22 thre -way junction. Th core of S17 is known from NMR t b a β-barr l with an OB fold, with long extended loops [23]. Thes loops are disord r d in solution but bind RNA in th 30S. In Th rmus, ther is a long C-terminal extension to S17 that is organized as an RNA-binding helix. The core of the protein and the C-terminal helix make extensive contacts with H11 and also contact H7. The C-terminal helix also contacts H21 in the central domain. Two long loops, loop 1 (26-36) and loop 2 (60-71) are ordered and interact with disparate domains of RNA exactly as predicted. Loop 1, which contains the site of neamine resistance, is inserted between H21 and a highly irregular structure at the base of H11. The very tip of loop 1 also touches the 560 loop of 16S RNA. Loop 2, which contains the site of a mutant defective in assembly, is involved in stitching together H7 and H11. Thus S17 interacts with H7, H11 and the 560 loop in the 5' domain, and H21 in the central domain.

[0090] **S16:** For a small protein, S16 has an extensive footprint throughout the 5' domain. All of the residues (1-88) are visible in the electron density, and were rebuilt using an NMR structure [24] as a guide. The protein consists of an N-terminal sheet with two extended loops, and two short helices in the C-terminal end. All of the extensive contacts with 16S RNA are now clear. The β-sheet is packed between the 608/620 internal loop of H21 on one side and a minor groove of H4 on the other. The two loops that extend out from this sheet both interact with RNA. Loop 1 interacts with phosphates in major groove of H4, while residues 39-43 in loop 2 make contact with the phosphate backbone around the internal loop near 453 in H17. The first helix (53-61) also extends across the major groove of this internal loop, while the C-terminal end of the second helix along with the turn leading out of it point into a minor groove of H17. There is also interaction with the 110 loop of the 5' domain. Finally, the extended C-terminus lies across the minor groove at the tip of H17.

[0091] **S20:** The current high resolution structure of S20 shows that the long N-terminal helix contacts the base of H6 and the tip of helix 44, and many conserved basic residues make salt-bridges with phosphates. Helices 2 and 3 of S20 interact with the minor groove of H9, and helix 3 also interacts with tip of H11 (263). Finally the extreme C-terminus of the protein is extended and lies along the minor groove of H9, which is longer in Thermus by 11 nucleotides. Thus S20 brings together several helices near the bottom of the subunit.

Proteins near the functional centre.

[0092] S4, S5 and S12 are clustered near the "functional center" of the ribosome and contain the sites of several important mutations.

[0093] **S4:** In the structure of isolated S4 [25, 26] the N-terminal domain was cleaved off prior to crystallization. This N-terminal region is organized as a tightly folded domain with a metal ion (presumably zinc) that is coordinated by four cysteines. The domain is packed against the body of the protein. While the N-terminus of S4 is highly conserved, the cysteines themselves are not. It is therefore likely that the addition of a "zinc finger" is for additional stability rather than essential for the fold. The linker residues 46-52 connect the N-terminal domain with the rest of the protein. All domains of S4 make intimate contacts with RNA. In particular, S4 makes extensive contacts with a five-way junction where H3, H4, H16, H17 and H18 come together in the 5' domain.

[0094] The N-terminal domain is packed against the 420 stem-loop (H16). The largely helical domain I is packed against a complicated region of RNA where H3 and the 507 bulge at the base of H18 come together. The remaining domain of S4 makes extensive contact with the minor groove of the base of H16. In addition, it also makes contact with the tip of the H21, which is itself packed against H4. This position is consistent with the large body of biochemical data on S4 binding to 16S RNA.

[0095] The C-terminus of S4 makes an extensive interface with S5. Most of the known mutations of S4 and S5 that confer the ram phenotype are located in this region [27, 28]. The interface consists of several highly conserved salt bridges, and some of the mutations break one or more of these interactions.

[0096] **S5:** The structure of S5 shows that the loop from residues 14-28 is folded back onto the body of the protein in the isolated structure, but is a fully extended β-hairpin in the 30S. Also, the C-terminus of S5, which is disordered in the isolated structure, is mainly helical and packs against a complicated surface of S8 formed by many different strands.

[0097] S5 interacts closely with a region of the ribosome where the head and the body come together. In the head, the extended H35/H36 helix packs against H28, which forms the neck of the 30S connecting the body with the head. The tip of H36 makes contact with H26a, H2 and the central pseudoknot in the body. Protein S5 has contacts throughout this region, ther by stabilizing th conf rmation of th h ad with respect to th body.

[0098] Th C-terminal sh et of S5 makes xtensiv int ractions with the major groov of th H1 and the c ntral pseudoknot. Th N-terminal domain binds to th major groove of H36, as does th bas of th β-hairpin. The tip of the hairpin interacts with th phosphat backbone in H28 and is also very close to H34. Nucl otide 560 is v ry close to

K121 in agre ment with crosslinking data.

[0099] Most of the xtensiv interactions with RNA occur via major grooves or phosphat backbon .

[0100] **S12:** S12 is unusual both for its structur and location. It is uniqu among th 30S prot ins in being on the interface side of the subunit. Its central core consists of a b-barr l with an OB fold, a f ature found in other proteins such as S17. This core binds tog ther H18, th 530 stem loop (at th tip of H18), H3 and a part of H44 clos to th decoding site. An unusual feature is a long extension that connects this core with a short helix at the N-terminus of the protein. This extension threads between the 560 loop and H12 on one side, and H11 on the other, to make contact with both S8 and S17 on the other side of the 30S.

[0101] S12 is also the only protein in the vicinity of the decoding site near 1492-1493 of RNA. It is the site of a number of functionally interesting mutations.

The head proteins S7 and S9.

[0102] **S7:** Protein S7, whose structure in isolation was previously known, is known to be crucial for the assembly of the head [29]. In our 30S structure, the entire sequence is visible, including the very basic N-terminus. S7 binds to a small but complex region of the tpd that encompasses two multiple-stem junctions at a corner of the head. The majority of the interaction surface consists of H29 tightly docked to the S-turn at the base of H43. This docking requires a tight turn at 1346, probably stabilized by S7 binding. Because S7 also makes interactions with H28, its primary surface of interaction encompasses all three of the helices around the H28/H29/H43 three-way junction. The very tight docking of H29 to H43 gives rise to a small region of very high negative charge density, which is bound by a surface of S7 with very high concentration of positive charge (mainly S7 helices 1 and 4).

[0103] The second important interaction surface is centred on the second multiple stem junction that S7 binds, the H29/H30/H41/H42 junction. In this junction, H30 and the base of H42 are tightly packed together, with a tight turn between them. An S-turn between helices 41 and 42 mediates packing of H41 and H42, which also have a tight turn between them. H41 also packs very tightly against H43. S7 makes contacts to the phosphate backbone of H41, stabilizing its packing with H43, and to residues around 1240 and 1298 where the tight bends occur in the H29/H30/H41/H42 junction. Contacts with U1240 are particularly intimate: the universally conserved bulge U1240 is deeply buried into a conserved hydrophobic pocket between the 35 and 115 loops of S7.

[0104] The β-hairpin is not tightly associated with 16S RNA, but probably fits tightly into the minor groove of the E-site tRNA. The structure is in rough agreement with a model of S7 binding to ribosomal RNA [30], but there are also significant differences, including the location of H43.

[0105] **S9:** S9 consists of a compact RNA-binding domain consisting of 2 helices packed against a 5-stranded sheet, with a third short helix at the C-terminal end of the domain. From this domain, there is a long 25 residue C-terminal tail that snakes into elements of the head RNA. S9 also interacts with S7 via a small hydrophobic patch.

[0106] The sheet of S9 makes extensive interactions with H38 and H39. It also has two loops that interact with the 1250 internal loop of H41. The short C-terminal helix interacts with 1177-1180 in H40.

[0107] The long C-terminal extension snakes between the H29-H43 junction on one side and the H38-H34 junction on the other to touch a portion of H31.

The S3 S10 S14 cluster.

[0108] These three proteins form a cluster on the rear left-hand of the head, as the protein portion of the beak. S3 is clearly stacked on top of the other two proteins, consistent with the order of assembly.

[0109] **S14:** S14 is bound in a crevice in the RNA and is mostly covered by S3 and S10. Almost the whole molecule contacts RNA, including helices H31, H32, H34, H38, and H43. A cross linked residue is in close proximity to the RNA 28.

[0110] S14 contains a zinc ion coordinated by four cysteines from a CXXC-X12-CXXC motif. This motif is structurally similar to that found in the first zinc finger in the glucocorticoid receptor. This zinc binding motif is not conserved among all bacteria, although many of the residues surrounding it are, suggesting perhaps that in other organisms the protein folds via a hydrophobic core.

[0111] **S10:** S10 is structurally very similar to the S6 fold, with two helices packed against a 4-stranded sheet. Two of the strands in this sheet are connected by a long β-hairpin that extends out from the sheet and is inserted right into the centre of the head RNA fold. The β-hairpin makes most of the contacts with RNA, including helices H31, H34 and H41. The two strands of the sheet pack into the shallow minor groove of H39, making contacts with backbone residues on both sides of the groove.

[0112] **S3:** S3 contains two domains, both consisting of two helices pack d against a 4-strand d sh et, which is similar to sev ral oth r ribosomal prot ins. In addition to the domains th re is an N-terminal tail (all of which is visibl). Th C-terminal 30 residues are poorly conserved and disordered in th structure.

[0113] RNA contact is mad by the N-terminal tail and th C-terminal domain. The N-terminal tail fits into a major

groove of H34. The short turn in the C-terminal domain also packs against H34.

[0114] The N-terminal domain makes few if any contacts with the RNA, but is mainly involved with making protein contacts with S10 and S14.

S13 and S19.

[0115] S13 and S19 form a loose dimer at the very "top" of the interface side of the head, extending both above and closer to the 50S than any of the head RNA. In spite of their location in this flexible region, they are both relatively well-defined in the electron density. Except for the C-terminal tail of S13, which reaches into the head and almost touches the tail of S9, none of these proteins are in contact with any other of the proteins in the small subunit. Together with S12, S11 and S15, these are among the few proteins that surround the region of intersubunit contact.

[0116] **S13:** All 125 residues of S13 are visible in the structure. The N-terminus (about 60 residues) forms a compact domain consisting of three small helices. Of this domain, only a small loop is in contact with the RNA and the domain appears to be clinging to the subunit only by virtue of its highly extended C-terminal region. This region begins with a long, straight alpha-helix that creeps along the top of the 30S head towards S19. It interacts mainly with the 1300 loop and H42. At this point the polypeptide chain bends by about 90 degrees, and the rest of the protein is mostly lacking in any secondary structure. This extended region curves around H41 into the head where it is buried in the RNA about 50-60 Å from the globular, N-terminal domain. It contacts H30 in the head.

[0117] **S19:** S19 consists of 92 residues. An NMR structure of isolated S19 [31] showed a single globular domain consisting of a helix packed against a three-stranded sheet, in which residues 9-78 were ordered. In the 30S structure, residues 2-81 are visible in the electron density. The C-terminus of the protein points towards the interface side and may become ordered in the 70S complex. Like S13, most of the globular domain of S19 is well separated from the RNA, but here both the N- and C-terminal extensions to the globular domain, as well as the loops 68-73 and 34-39 make contacts with H42. The C-terminal extension, like S13, bends around the RNA, to contact H31 while the N-terminus reaches H42 some considerable distance away. Thus, S19 straddles a portion of the head of the 30S. The residues in S13 and S19 that were crosslinked 48 are adjacent to each other in the structure.

S2.

[0118] Thermus S2 consists of 256 residues of which 7-235 are visible in our structure. The protein consists of a large central domain of about 200 residues that consists of a 5-stranded parallel sheet and four helices connecting them. Two helices that form a small coiled-coil motif protrude out of this domain. The protein is located on the back of the 30S at the interface between the head and the rest of the particle. While it is primarily regarded as a "head" protein, it also makes contacts with the central domain in our structure.

Thx.

[0119] This small 26 residue peptide was isolated and characterized from Thermus ribosomes [18]. Thx fills a cavity formed by a number of different elements at the very top of the head. Residues 1-24 are visible in the electron density, of which 8-14 form a short helix, flanked by extended ends. It is surrounded by H42, the tip of H41, and the base of H41, while the bottom of the cavity is formed by the major groove of H43. The protein is highly basic, and there are extensive salt-bridges between these residues and phosphates of nearby RNA. Thus Thx stabilizes a number of different RNA elements that come close together near the top of the head.

FUNCTIONAL INSIGHTS FROM THE STRUCTURE OF THE 30S RIBOSOMAL SUBUNIT.

[0120] During translation of the genetic code, the 30S ribosomal subunit provides the framework for base-pairing between the anticodon of tRNA and the codon of mRNA, and discriminates between cognate and non-cognate tRNAs to ensure translational fidelity, in a process termed decoding. During translocation, the ribosome must move by precisely one codon relative to mRNA and the bound tRNAs. Both decoding and translocation involve "switches" in which precise conformational changes occur in the ribosome. The atomic resolution structure of the 30S subunit allows us to interpret the environment of the mRNA and tRNA binding sites in molecular terms. In one well-characterized example of a functional switch involved in accuracy, we are also able to determine the spatial arrangement of its elements, thus elucidating its architecture. The structure also suggests other possible switching elements in the 30S, and sheds light on the kinds of movements that might occur.

[0121] The ribosome contains three tRNA binding sites, designated A (aminoacyl), P (peptidyl) and E (exit), after their respective tRNA substrates. Each site is bipartite, located partly on the 30S ribosomal subunit and partly on the 50S subunit. The A- and P-site tRNAs bind with their aminoacyl acceptor ends on the 50S subunit, and with their

anticodon ends base-paired to adjac nt mRNA codons on th 30S subunit. The E-sit tRNA is bound in a similar orientation but it is not known wheth r th E-site tRNA is base-paired to th E-site mRNA codon. The 30S subunit also binds mRNA upstream and downstream of th A, P and E codons. During translation, incoming aminoacyl tRNA is delivered to the A-site as a ternary compl x with EF-Tu and GTP. Discrimination of cognate from non-cognate tRNAs occurs in th A-sit . It is thought that th re is also a second "proofreading" discrimination st p in th A-sit after GTP hydrolysis by EF-Tu, which is needed to discriminate cognate from near-cognate tRNAs. The 30S P-site has a much higher affinity for tRNA, in order to maintain the reading frame.

[0122] There is one well-characterized conformational switch in the 30S subunit, the helix 27 accuracy switch [32]. Genetic and biochemical data support a model in which this switch may be part of a larger-scale conformational change that occurs between initial selection and proofreading of the A-site tRNA, or the switch may play a role in translocation.

[0123] Until recently, there has been a large disparity between the high resolution of the genetic and biochemical data that define the RNA components of the active sites of the 30S subunit, and the relatively low-resolution of the three-dimensional structures of ribosomes available. The present invention addresses this disparity. In combination with previous biochemical and other data, it is now possible to identify the detailed structure of 30S active sites. In addition, by superimposing the tRNA and mRNA coordinates from the known 7.8 Å 70S structure, it is now possible to infer many of the interactions between 30S active sites and tRNA/mRNA ligands.

[0124] With our complete and high resolution structure of the 30S subunit in hand, it is now possible to identify at the residue level the elements of the 30S subunit that interact with the anticodon stem-loop (ASL) of the A, P and E-site tRNAs and associated mRNA.

[0125] Identification of the precise boundaries of the A, P, and E sites in an unbiased fashion in a structure determined in the absence of cognate tRNA ligands would normally be problematic. As it happens, the P-site in the 30S structure is filled with a stem-loop of RNA corresponding to residues 75-95 (in the *E. coli* numbering system) from the end of the "spur" (H6) of a neighbouring molecule. (Henceforth the term "spur" will refer to the symmetry-related spur docked in the P-site, rather than the spur at the bottom of the same subunit). The spur appears to mimic P-site tRNA by a variety of criteria. The extent of the 30S interaction with the anticodon stem-loop (ASL) is in very good agreement with that determined by affinity measurements [33] and by hydroxyl radical footprinting [34]. Secondly, the conformation of the spur stem-loop is distorted in order to more closely resemble the canonical tRNA ASL conformation [35, 36]: a U-A base pair is broken so that the spur hairpin loop can approximate the conformation of a tRNA ASL, complete with a U-turn and stacked anticodon. Another indication that the spur is a mimic of a bound P-site tRNA ASL is that of the twelve hydrogen bonds between 30S and the spur, only one appears to be sequence-specific, in accordance with the lack of sequence conservation in tRNA anticodon stems. Finally, close contacts of the spur with 16S RNA are on the whole consistent with chemical protection data for P-site tRNA [37] and with the 34-C1400 UV-induced crosslink between tRNA and 16S RNA [38] (the analogous residues are stacked in the 30S crystal structure).

[0126] Yet another indication that the spur mimics a P-site tRNA ASL is that its "pseudo-anticodon" is base-paired to a triplet of nucleotides, a mimic of mRNA. A fourth nucleotide is also visible 5' to the pseudo-anticodon, in the E site. These pseudo-codon bases are clearly pyrimidines, and appear to be UCU from the base-pairing geometries, which are U-U, U-C, and U-U since the pseudo-anticodon is UUU. The origin of this "pseudo-message" is unclear, but it probably comes from the 3' end of 16S RNA, which ends with 5' U1542C1543U1544 3'. The last nucleotide of our 16S model is C1533, so that seven disordered nucleotides would span the 25Å gap between C1533 and U1541, which is clearly stereochemically feasible. Alternatively, it is possible that the 3' end of 16S RNA has been cleaved somewhere between C1533 and U1541 prior to or during crystallization. The presence of functional mimics of mRNA and P-site tRNA also explains why these crystals diffract relatively well: the P-site tRNA makes extensive contacts with both the head and the body of the 30S, thereby helping to lock the particle into a single conformation.

[0127] To ask how well pseudomessage and spur mimic mRNA and the ASL of tRNA, we have used the 7.8 Å resolution structure of the 70S ribosome with bound mRNA and tRNAs [39]. In that structure, two elements of 16S RNA were identified, H27 and H44. To avoid any possible bias in our interpretation of the spur as a mimic, only H27 and H44 were used to in the alignment to superimpose the 70S structures onto our 30S structure. Despite the relatively low resolution of the 70S structure used, a least-squares superposition of these two elements had a phosphate r.m.s. d. of only 2.3 Å. When the 70S elements are superimposed in this manner onto our 30S structure, we found that indeed, as expected, the P-site tRNA superimposes well onto the 30S spur, and the 30S pseudo-message corresponds to the P-site codon. In particular, the orientation of the spur stem-loop is very similar to the 70S P-site ASL, and there are no significant clashes between the 70S A- and E-site tRNAs and our 30S subunit when superimposed in this manner. It is clear that the spur and pseudo-message cannot be perfect mimics, however, because the pseudo-anticodon - codon h lix consists of thre pyrimidine-pyrimidin base pairs, which are about 2 Å narrower than Watson-Crick pairs. Thus it s ems likely that the spur and its ps udo-messag are good but not perfect mimics of P-site tRNA and P-sit codon, respectively, and that th spur mimic mod l should xplain many but perhaps not all f atur s of P-sit tRNA binding to the 30S. Moreover, th transform d A- and P-site tRNAs and A-site codon provides a useful landmarks for modeling th extent of the A- and E-sites of th 30S.

The P-site.

[0128] The P-site spur contacts several discrete regions of 16S RNA, most of which have been implicated in P-site binding by biochemical experiments. Two proteins also participate in binding the P-site ASL, a possibly surprising result. Most of the contact surface lies between the minor groove of the spur stem and 16S RNA nucleotides 1338-1341, 1229-1230, and the C-terminal tails of proteins S13 and S9. There are many hydrogen bonds between the minor groove (i.e. the 2' OH and base groups) of spur residues C91, C92, and G78 and the minor groove surface of G1338-A1339. Only one of these hydrogen bonds appears to be sequence-specific (G78 N2 - A1339 N3). A contact from Lys 126 of S9 appears to help stabilize this minor-groove to minor-groove packing interaction. Both 1338 and 1339 have previously been implicated in P-site binding [37]. A second area of contact, nearly continuous with the first, is between the 16S 1229-1230 sugar-phosphate backbone and spur residues G77 and G78. This region of contact is extended by the C-terminal tail of S13, which seems to help glue the spur and the 1229-1230 area together. The other areas of contact are much more tenuous. One interaction is stacking of U82 on C1400, which rationalizes the ASL 34-C1400 uv-induced crosslink [38]. The other is a packing interaction between A790 and spur residues 88-89, with a single hydrogen bond present. A790 is a so-called class III site, that is it is protected by either tRNA or 50S subunits. From the spur interaction, it would thus appear that binding of either the 50S subunit or the P-site ASL stabilizes a contact between A790 N6 and the phosphate of 1498, i.e. a contact between the central and three-prime minor domains. Finally, if the pseudo-codon - pseudo-anticodon helix were a few Å wider, as it would be for a Watson-Crick-paired helix, it would make van der Waals contact with the base of G966. G966 has also previously been implicated as part of the P-site by chemical modification experiments and has also been identified as a one of the few guanines crucial for P-site binding [40].

[0129] The P-site codon is threaded through the major groove of the upper portion of helix 44, in a universally conserved region of 16S RNA. There appears to be a tight turn between nucleotides -1 and +1, that is, between the last E-site and the first P-site codon nucleotides. This tight turn is stabilized by a hydrogen bond to the N1/N2 groups of the conserved residue G926, a residue previously implicated as crucial for P-site binding [40]. Additional hydrogen bonds are seen between the 2'OH of +1 to the phosphate of C1498, and between the phosphate of +2 and the 2' OH of C1498. The phosphate of +2 also stacks on the base of C1498. The phosphate of +3 is within hydrogen-bonding distance of two conserved cytidine N4 groups, from C1402 and C1403. The +3 base also stacks on the sugar of C1400. Finally, it appears likely that there are several magnesium ions that may help stabilize the location of the P-site codon in the major groove of H44.

The E-site.

[0130] The E-site is defined by the environment surrounding the 70S E-site tRNA superimposed onto our 30S structure, as described above. Unlike the A and P-sites, the E-site consists mostly of protein. Proteins S7 and S11 have a small interface that binds the minor groove of the E-site ASL. The highly conserved beta-hairpin of S7 extends this surface nearly to the bottom of the anticodon, and it is possible that the S7 beta-hairpin helps dissociate the E-site codon from the E-site anticodon. The RNA portion of the E-site makes only tenuous interactions with the E-site ASL. 16S nucleotides 1382 and 1383 may interact with residue 34 of the anticodon. The minor-groove surface of the conserved 16S residues 693 and 694 may interact with the minor-groove surface of the 37-39 residues of the E-site ASL.

The A-site.

[0131] The A-site is rather wider and shallower than the P or E sites, perhaps in order to allow rotation of the A-site codon-anticodon helix during or after GTP hydrolysis by EF-Tu. The RNA components of the A-site appear to include portions of the 530 loop, H34 in the head, and residues 1492-1493 from the 3' minor domain, all of which have been previously implicated in A-site binding.

The Helix 27 switch.

[0132] It is clear that many of the elements that make contact with the various tRNA would have to move during translocation. Indeed, the ribosome is known to undergo extensive conformational changes during the elongation cycle, and these must involve breaking and making precise contacts. However, the precise switching elements in these conformational changes are not known in general, with the exception of a switch in H27.

[0133] H27 is proposed to have two alternative base-pairing schemes during translation, one a *"ram"* or permissive form that pairs 885-887 with 910-912, and an alternative "restrictive" form that pairs 888-890 with 910-912 [32]. The *ram* form appears to be the more stable form in the ribosome and it features an S-turn (or loop E motif) in H27. The S-turn in H27 is also seen in the tRNA-bound structure of the 70S [39]. A switch to the restrictive form would involve a sliding of the two strands of H27 relative to each other and the S-turn would be replaced by an internal loop with a

different structure for H27. Inde d, analysis of the two forms by cryoelectron microscopy r veal noticeable conformational changes in th ribosom , specially around th A-site [41]. W can now precisely define th structure around H27 and use previous chemical modification data [32] to suggest th kinds of mov ment involv d.

[0134] The S-turn in H27 around 888 is right next to 1489 in H44, and H27 packs against the minor groov of H44 just b low the decoding sit . Th tip of H27 is close to H11, while 885, which is bas -paired with 910 in our conformation, is near both H1 and the 570 loop. Finally, 914 is near both H1 and 526 in the 530 loop. Thus H27 is right in the heart of an area which includes the decoding site and the 530 loop. So it is not surprising that a change in the conformation of H27 would have affect these elements.

[0135] A number of elements that are more accessible in the "restrictive" state appear to be protected in the structure of the present invention. Thus for example, 524-526 are currently base-paired with 507-505 in the 530 pseudoknot. This suggests that the 530 pseudoknot could be broken in the restrictive state. Similarly, 1053 and 1197 are base-paired in the current structure, but they are part of a distorted region of H34 analogous to an S-turn, and it is not hard to envisage that an analogous switch might occur in H34 in the alternative state. Thus the data in combination with our structure suggests that H34 in the head and the 530 loop in the shoulder move between the two states, with H34 possibly adopting an alternative form, and the 530 pseudoknot being disrupted. In this context, it is interesting to note that both H34 and the 530 loop have been implicated in tRNA binding.

[0136] Other parts of the chemical protection data, especially those that are supposed to indicate enhanced accessibility in the *ram* state, are not so easy to rationalize since they involve protected bases in our structure.

[0137] The 30S structure has allowed us to identify details of the tRNA and mRNA binding sites, as well as obtain our first detailed look at the structure around the H27 switch. Clearly, H27 is only one component of major conformational changes that occur during translation. Analysis of the high resolution 30S structure should allow us to identify other potential switching elements, which may then be tested genetically.

References:

[0138]

1. Garrett, R. A. et al. (eds.) The Ribosome. Structure, Function, Antibiotics and Cellular Interactions (ASM Press, Washington, D.C., 2000).

2. von Böhlen, K. et al. Characterization and preliminary attempts for derivatization of crystals of large ribosomal subunits from Haloarcula marismortui diffracting to 3 Å resolution. J. Mol. Biol. 222, 11-15 (1991).

3. Trakhanov, S. D. et al. Crystallization of 70 S ribosomes and 30 S ribosomal subunits from Thermus thermophilus. FEBS Lett. 220, 319-322 (1987).

4. Glotz, C. et al. Three-dimensional crystals of ribosomes and their subunits from eu- and archaebacteria. Biochem. Int. 15, 953-960 (1987).

5. Yonath, A. et al. Characterization of crystals of small ribosomal subunits. J. Mol. Biol. 203, 831-834 (1988).

6. Yusupov, M. M., Tischenko, S. V., Trakhanov, S. D., Ryazantsev, S. N. & Garber, M. B. A new crystalline form of 30 S ribosomal subunits from Thermus thermophilus. FEBS Lett. 238, 113-115 (1988).

7. Yonath, A. et al. Crystallographic studies on the ribosome, a large macromolecular assembly exhibiting severe nonisomorphism, extreme beam sensitivity and no internal symmetry. Acta Crystallogr A54, 945-55 (1998).

8. Tocilj, A. et al. The small ribosomal subunit from Thermus thermophilus at 4.5 A resolution: pattern fittings and the identification of a functional site. Proc Natl Acad Sci U S A 96, 14252-7 (1999).

9. Clemons, W. M., Jr. et al. Structure of a bacterial 30S ribosomal subunit at 5.5 Å resolution. Nature 400, 833-840 (1999).

10. Otwinowski, Z. & Minor, W. in Methods in Enzymology (eds. Carter, C. W. J. & Sweet, R. M.) 307-25 (Academic Pr ss, N w York, 1997).

11. T rwillig r, T. & Ber ndz n, J. Automated MAD and MIR structur d termination. Acta Cryst D55, 849-861 (1999).

12. Abrahams, J. P. Bias reduction in phas refinement by modified int rference functions: introducing the gamma correction. Acta Cryst. D53 (1997).

13. d la Fortelle, E. & Bricogne, G. in M thods in Enzymology (eds. Carter, C. W., Jr. & Sweet, R. M.) 472-93 (Academic Press, New York, 1997).

14. Hartmann, R. K. & Erdmann, V. A. Thermus thermophilus 16S rRNA is transcribed from an isolated transcription unit. J Bacteriol 171, 2933-41 (1989).

15. Cowtan, K. & Main, P. Miscellaneous algorithms for density modification. Acta Crystallogr D Biol Crystallogr 54, 487-93 (1998).

16. Jones, T. A. & Kjeldgaard, M. Electron-density map interpretation. Meth. Enzymol. 277B, 173-207 (1997).

17. Brünger, A. T. et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 54, 905-21 (1998).

18. Choli, T., Franceschi, F., Yonath, A. & Wittmann-Liebold, B. Isolation and characterization of a new ribosomal protein from the thermophilic eubacteria, Thermus thermophilus, T. aquaticus and T. flavus. Biol Chem Hoppe Seyler 374, 377-83 (1993).

19. Mueller, F. & Brimacombe, R. A new model for the three-dimensional folding of Escherichia coli 16 S ribosomal RNA. I. Fitting the RNA to a 3D electron microscopic map at 20 A. J Mol Biol 271, 524-44 (1997).

20. Mougel, M. et al. Minimal 16S rRNA binding site and role of conserved nucleotides in Escherichia coli ribosomal protein S8 recognition. Eur J Biochem 215, 787-92 (1993).

21. Wu, H., Jiang, L. & Zimmermann, R. A. The binding site for ribosomal protein S8 in 16S rRNA and spc mRNA from Escherichia coli: minimum structural requirements and the effects of single bulged bases on S8-RNA interaction. Nucleic Acids Res 22, 1687-95 (1994).

22. Urlaub, H., Thiede, B., Muller, E. C., Brimacombe, R. & Wittmann-Liebold, B. Identification and sequence analysis of contact sites between ribosomal proteins and rRNA in Escherichia coli 30 S subunits by a new approach using matrix-assisted laser desorption/ionization-mass spectrometry combined with N-terminal microsequencing. J Biol Chem 272, 14547-55 (1997).

23. Golden, B. L., Hoffman, D. W., Ramakrishnan, V. & White, S. W. Ribosomal protein S17: characterization of the three-dimensional structure by 1H- and 15N-NMR. Biochemistry 32, 12812-20 (1993).

24. Allard, P. et al. Another piece of the ribosome: Solution structure of S16 and its location in the 30S subunit. Structure, (2000).

25. Davies, C., Gerstner, R. B., Draper, D. E., Ramakrishnan, V. & White, S. W. The crystal structure of ribosomal protein S4 reveals a two-domain molecule with an extensive RNA-binding surface: one domain shows structural homology to the ETS DNA-binding motif. Embo J 17, 4545-58 (1998).

26. Markus, M. A., Gerstner, R. B., Draper, D. E. & Torchia, D. A. The solution structure of ribosomal protein S4 delta41 reveals two subdomains and a positively charged surface that may interact with RNA. Embo J 17, 4559-71 (1998).

27. van Acken, U. Proteinchemical studies on ribosomal proteins S4 and S12 from ram (ribosomal ambiguity) mutants of Escherichia coli. Mol Gen Genet 140, 61-8 (1975).

28. Wittmann-Liebold, B. & Greuer, B. The primary structure of protein S5 from the small subunit of the Escherichia coli ribosome. FEBS Lett 95, 91-8 (1978).

29. Nowotny, V. & Ni rhaus, K. H. Assembly of th 30S subunit from Escherichia coli ribosomes occurs via two assembly domains which ar initiated by S4 and S7. Biochemistry 27, 7051-5 (1988).

30. Tanaka, I. t al. Matching th crystallographic structur of ribosomal prot in 37 to a thre -dimensional model of th 16S ribosomal RNA. Rna 4, 542-50 (1998)

31. Helgstrand, M. et al. Solution structure of the ribosomal prot in S19 from Thermus thermophilus. J Mol Biol 292, 1071-81 (1999).

32. Lodmell, J. S. & Dahlberg, A. E. A conformational switch in Escherichia coli 16S ribosomal RNA during decoding of messenger RNA. *Science* **277,** 1262-7 (1997).

33. Rose, S. J. d., Lowary, P. T. & Uhlenbeck, O. C. Binding of yeast tRNAPhe anticodon arm to Escherichia coli 30 S ribosomes. *J Mol Biol* **167,** 103-17 (1983).

34. Huttenhofer, A. & Noller, H. F. Hydroxyl radical cleavage of tRNA in the ribosomal P-site. *Proc Natl Acad Sci U S A* **89,** 7851-5 (1992).

35. Jack, A., Ladner, J. E. & Klug, A. Crystallographic refinement of yeast phenylalanine transfer RNA at 2-5A resolution. *J Mol Biol* **108,** 619-49 (1976).

36. Rich, A. & RajBhandary, U. L. Transfer RNA: molecular structure, sequence, and properties. *Annu Rev Biochem* **45,** 805-60 (1976).

37. Moazed, D. & Noller, H. F. Binding of tRNA to the ribosomal A and P-sites protects two distinct sets of nucleotides in 16 S rRNA. *J Mol Biol* **211,** 135-45 (1990).

38. Prince, J. B., Taylor, B. H., Thurlow, D. L., Ofengand, J. & Zimmermann, R. A. Covalent crosslinking of tRNA1Val to 16S RNA at the ribosomal P-site: identification of crosslinked residues. *Proc Natl Acad Sci U S A* **79,** 5450-4 (1982).

39. Cate, J. H., Yusupov, M. M., Yusupova, G. Z., Earnest, T. N. & Noller, H. F. X-ray crystal structures of 70S ribosome functional complexes [see comments]. *Science* **285,** 2095-104 (1999).

40. von Ahsen, U. & Noller, H. F. Identification of bases in 16S rRNA essential for tRNA binding at the 30S ribosomal P-site. *Science* **267,** 234-7 (1995).

41. Gabashvili, I. S. *et al.* Major rearrangements in the 70S ribosomal 3D structure caused by a conformational switch in 16S ribosomal RNA. *Embo J* **18,** 6501-7 (1999).

TABLE 2

REMARK 465 MISSING RESIDUES
REMARK 465 THE FOLLOWING
RESIDUES WERE NOT LOCATED IN
THE
REMARK 465 EXPERIMENT.
(M=MODEL NUMBER; RES=RESIDUE
NAME; C=CHAIN
REMARK 465 IDENTIFIER;
SSSEQ=SEQUENCE NUMBER;
I=INSERTION CODE.)
REMARK 465

	M	RES	C	SSSEQI
REMARK 465		U	A	0
REMARK 465		U	A	1
REMARK 465		U	A	2
REMARK 465		G	A	3
REMARK 465		U	A	4
REMARK 465		C	A	1535
REMARK 465		C	A	1536
REMARK 465		U	A	1537
REMARK 465		C	A	1538
REMARK 465		C	A	1539
REMARK 465		U	A	1540
REMARK 465		U	A	1541
REMARK 465		U	A	1542
REMARK 465		C	A	1543
REMARK 465		U	A	1544
REMARK 465	MET		B	1
REMARK 465	PRO		B	2
REMARK 465	VAL		B	3
REMARK 465	GLU		B	4
REMARK 465	ILE		B	5
REMARK 465	THR		B	6
REMARK 465	GLU		B	241
REMARK 465	ALA		B	242
REMARK 465	GLU		B	243
REMARK 465	ALA		B	244
REMARK 465	THR		B	245
REMARK 465	GLU		B	246
REMARK 465	THR		B	247
REMARK 465	PRO		B	248
REMARK 465	GLU		B	249
REMARK 465	GLY		B	250
REMARK 465	GLU		B	251
REMARK 465	SER		B	252
REMARK 465	GLU		B	253
REMARK 465	VAL		B	254
REMARK 465	GLU		B	255
REMARK 465	ALA		B	256
REMARK 465	MET		C	1
REMARK 465	ILE		C	208
REMARK 465	GLY		C	209
REMARK 465	GLY		C	210
REMARK 465	GLN		C	211
REMARK 465	LYS		C	212
REMARK 465	PRO		C	213
REMARK 465	LYS		C	214
REMARK 465	ALA		C	215
REMARK 465	ARG		C	216
REMARK 465	PRO		C	217
REMARK 465	GLU		C	218
REMARK 465	LEU		C	219
REMARK 465	PRO		C	220
REMARK 465	LYS		C	221

REMARK 465	ALA C	222		REMARK 465	MET J	1	
REMARK 465	GLU C	223		REMARK 465	PRO J	2	
REMARK 465	GLU C	224		REMARK 465	VAL J	101	
REMARK 465	ARG C	225		REMARK 465	GLY J	102	
REMARK 465	PRO C	226		REMARK 465	GLY J	103	
REMARK 465	ARG C	227		REMARK 465	GLY J	104	
REMARK 465	ARG C	228		REMARK 465	ARG J	105	
REMARK 465	ARG C	229		REMARK 465	MET K	1	
REMARK 465	ARG C	230		REMARK 465	ALA K	2	
REMARK 465	PRO C	231		REMARK 465	LYS K	3	
REMARK 465	ALA C	232		REMARK 465	LYS K	4	
REMARK 465	VAL C	233		REMARK 465	PRO K	5	
REMARK 465	ARG C	234		REMARK 465	SER K	6	
REMARK 465	VAL C	235		REMARK 465	LYS K	7	
REMARK 465	LYS C	236		REMARK 465	LYS K	8	
REMARK 465	LYS C	237		REMARK 465	LYS K	9	
REMARK 465	GLU C	238		REMARK 465	VAL K	10	
REMARK 465	GLU C	239		REMARK 465	MET L	1	
REMARK 465	MET D	1		REMARK 465	VAL L	2	
REMARK 465	MET E	1		REMARK 465	ALA L	3	
REMARK 465	PRO E	2		REMARK 465	LEU L	4	
REMARK 465	GLU E	3		REMARK 465	ALA L	129	
REMARK 465	THR E	4		REMARK 465	LYS L	130	
REMARK 465	GLU E	155		REMARK 465	THR L	131	
REMARK 465	ALA E	156		REMARK 465	ALA L	132	
REMARK 465	HIS E	157		REMARK 465	ALA L	133	
REMARK 465	ALA E	158		REMARK 465	LYS L	134	
REMARK 465	GLN E	159		REMARK 465	LYS L	135	
REMARK 465	ALA E	160		REMARK 465	MET M	1	
REMARK 465	GLN E	161		REMARK 465	MET N	1	
REMARK 465	GLY E	162		REMARK 465	MET O	1	
REMARK 465	MET G	1		REMARK 465	ALA P	84	
REMARK 465	MET I	1		REMARK 465	ARG P	85	

REMARK 465 GLU P 86
REMARK 465 GLY P 87
REMARK 465 ALA P 88
REMARK 465 MET Q 1
REMARK 465 MET R 1
REMARK 465 SER R 2
REMARK 465 THR R 3
REMARK 465 LYS R 4
REMARK 465 ASN R 5
REMARK 465 ALA R 6
REMARK 465 LYS R 7
REMARK 465 PRO R 8
REMARK 465 LYS R 9
REMARK 465 LYS R 10
REMARK 465 GLU R 11
REMARK 465 ALA R 12
REMARK 465 GLN R 13
REMARK 465 ARG R 14
REMARK 465 ARG R 15
REMARK 465 MET S 1
REMARK 465 GLY S 82
REMARK 465 HIS S 83
REMARK 465 GLY S 84
REMARK 465 LYS S 85
REMARK 465 GLU S 86
REMARK 465 ALA S 87
REMARK 465 LYS S 88
REMARK 465 ALA S 89
REMARK 465 THR S 90
REMARK 465 LYS S 91
REMARK 465 LYS S 92
REMARK 465 LYS S 93
REMARK 465 MET T 1
REMARK 465 ALA T 2
REMARK 465 GLN T 3
REMARK 465 LYS T 4
REMARK 465 LYS T 5
REMARK 465 PRO T 6
REMARK 465 LYS T 7
REMARK 465 LYS V 26
REMARK 465 LYS V 27

Claims

1. A crystal of a 30S subunit having a tetragonal spac group $P4_12_12$ with unit cell dimensions of a = 401.375 Å, b = 401.375 Å, c = 175.887 Å.

2. A crystal of a 30S subunit having a tetragonal space group $P4_12_12$ with unit cell dimensions of a = 401.4 Å, b = 401.4 Å, c = 175.9 Å.

3. A crystal of a 30S ribosomal subunit having a resolution better (numerically less) than about 3 Å.

4. A crystal a 30S ribosomal subunit having the structure defined by the co-ordinates of Table 1.

5. A computer-based method of rational drug design which comprises:

 providing the structure of a 30S ribosomal subunit as defined by the coordinates of Table 1;
 providing the structure of a candidate modulator molecule; and
 fitting the structure of candidate to the structure of the 30S of Table 1.

6. A computer-based method for identifying a potential inhibitor of the 30S ribosome comprising the steps of:

 a. employing a three-dimensional structure of 30S, or at least one sub-domain thereof, to characterise at least one active site, the three-dimensional structure being defined by atomic coordinate data according to Table 1; and
 b. identifying the potential inhibitor by designing or selecting a compound for interaction with the active site.

7. The method of claim 6 which further comprises:

 c. obtaining or synthesising the potential inhibitor;
 d. contacting the potential inhibitor with 30S to determine the ability of said inhibitor to interact with the 30S.

8. The method of claim 6 which further comprises:

 c. obtaining or synthesising said potential ligand;
 d. forming a complex of 30S and said potential ligand; and
 e. analysing said complex by X-ray crystallography to determine the ability of said potential ligand to interact with 30S.

9. A method for the determination of the structure of a bacterial 30S from a species other than *T. thermophilus* which method comprises:

 (a) crystallising the 30S of said species to obtain a crystal;
 (b) performing X-ray crystallography on said crystal to obtain X-ray diffraction data;
 (c) providing the structure data of Table 1; and
 (d) using molecular replacement to calculate an electron density map of the 30S.

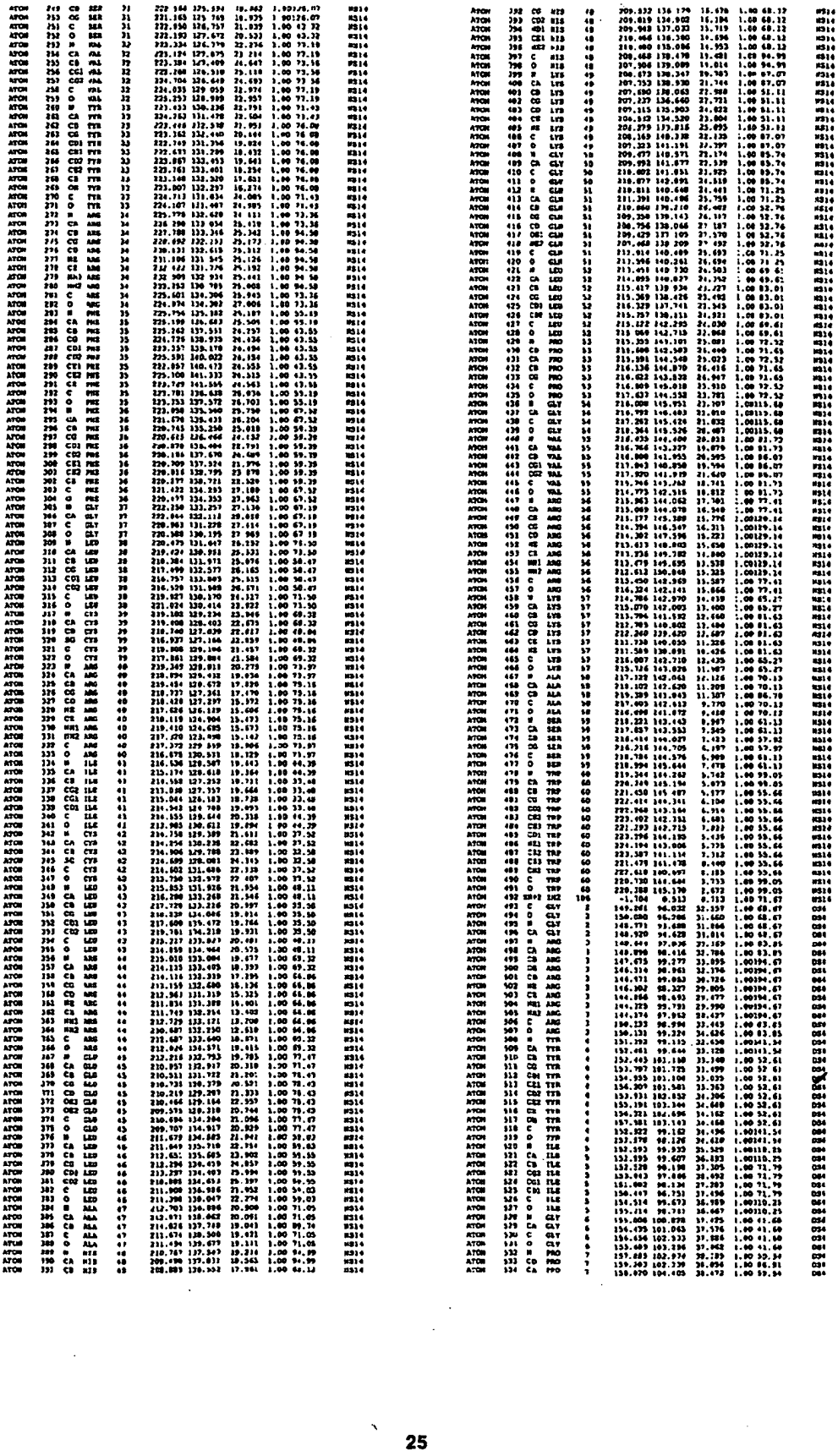
10. A computer system, intended to generate structures and/or perform rational drug design for the 30S ribosome or complexes of the 30S ribosome with a potential modulator, the system containing either (a) atomic coordinate data according to Table 1, said data defining the three-dimensional structure of 30S or at least one sub-domain thereof, or (b) structure factor data for 30S, said structure factor data being derivable from the atomic coordinate data of Table 1.

11. A computer readable media with either (a) atomic coordinate data according to Table 1 recorded thereon, said data defining the three-dimensional structure of the 30S ribosome, at least one atom or at least one sub-domain thereof, or (b) structure factor data for the 30S ribosome recorded thereon, the structure factor data being derivable from th atomic coordinat data of Tabl 1.

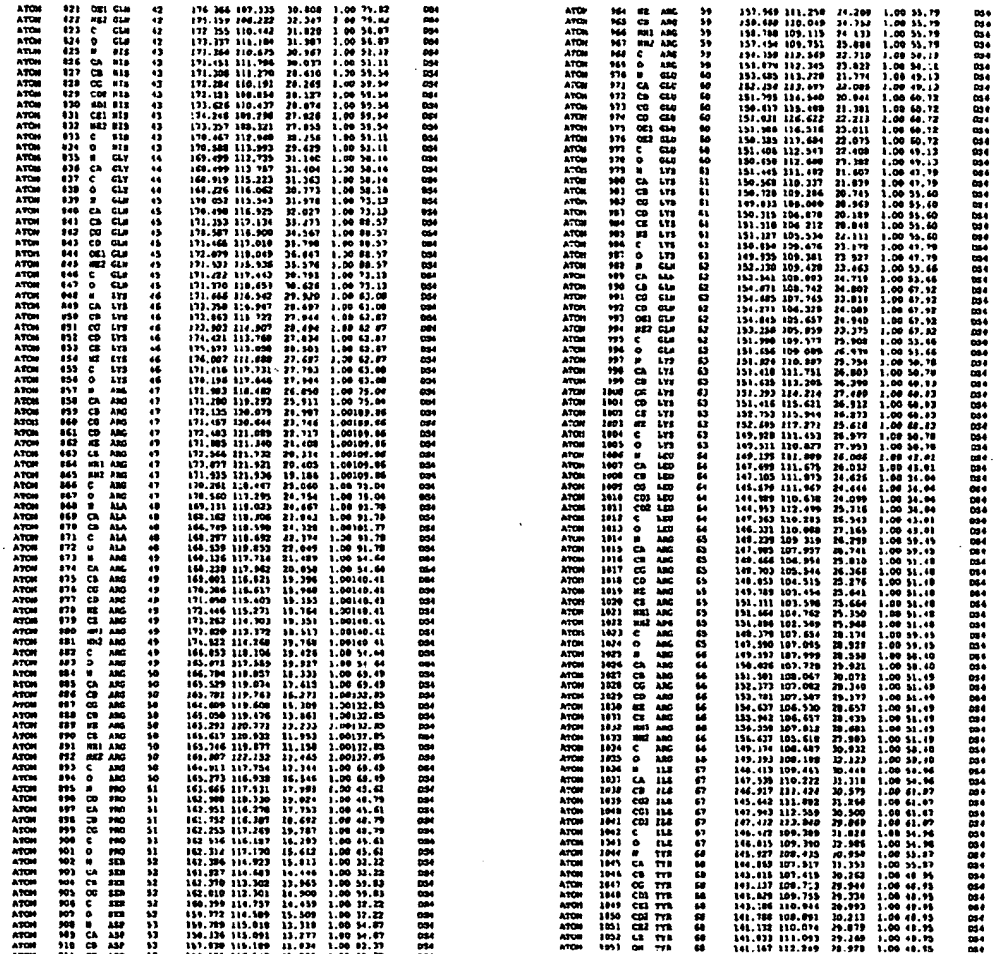
CRYSTAL STRUCTURE OF THE 30S RIBOSOME AND ITS USE - FIGURE 2

TABLE 1A

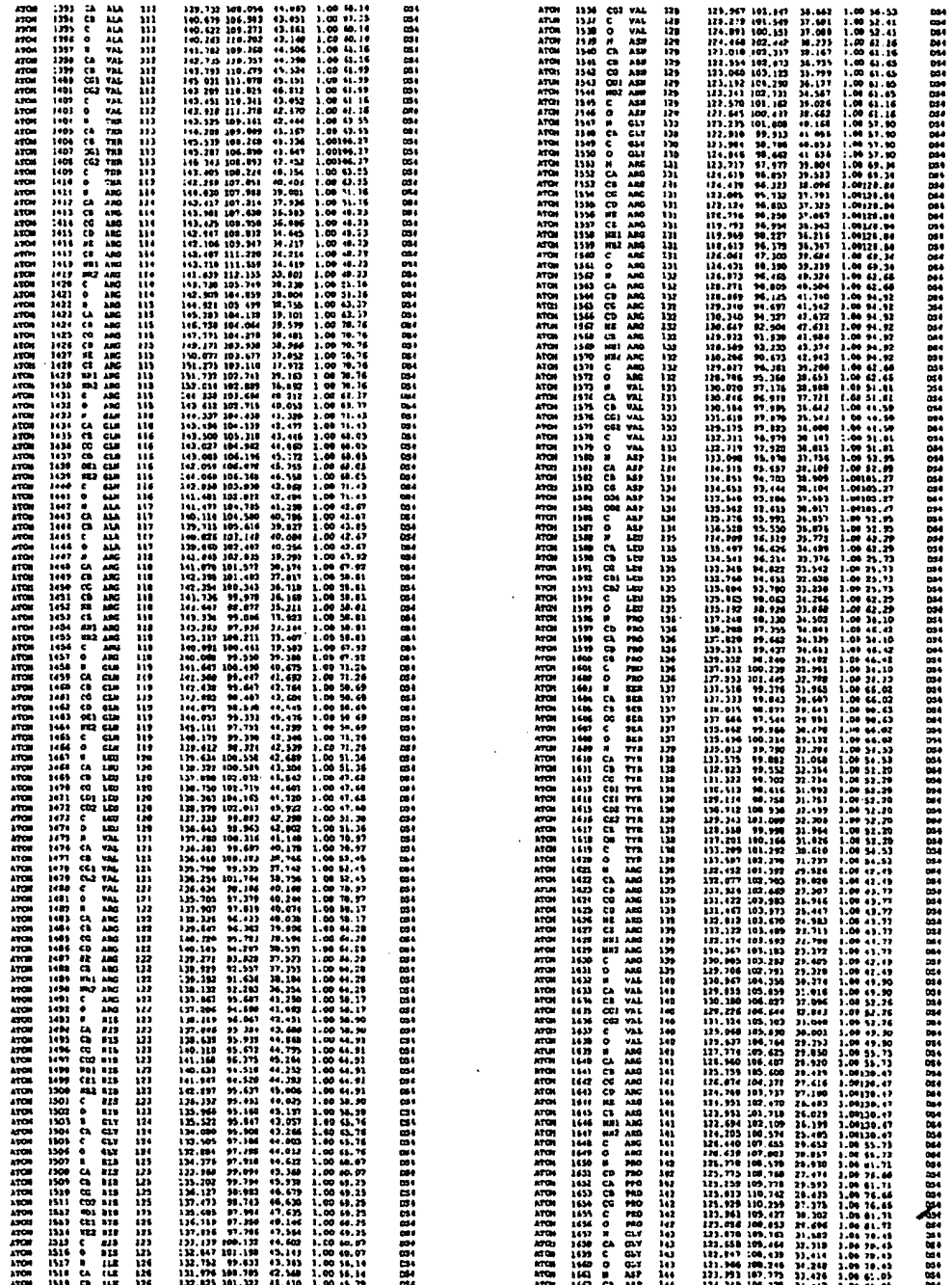

```
REMARK Coordinates from individual occupancy refinement
REMARK refinement resolution: 99 - 3.0 A
REMARK starting r= 0.2138 free_r= 0.2561
REMARK final   r= 0.2131 free_r= 0.2561
REMARK rms= 3.95364
REMARK targets als  steps= 10
REMARK sg= P6(1)22(154) a= 401.375 b= 401.375 c= 175.887 alpha= 90 beta= 90 gamma=
90
REMARK parameter file 1  : CNS_TOPPAR:protein_rep.param
REMARK parameter file 2  : CNS_TOPPAR:dna-rna-multi-endo.param
REMARK parameter file 3  : CNS_TOPPAR:ion.param
REMARK molecular structure file: generate.mtf
REMARK reset name
REMARK B-correction resolution: 6.0 - 3.0
REMARK initial B-factor correction applied to fobs :
REMARK     B11=   5.012 B22=   5.012 B33= -10.024
REMARK     B12=   0.000 B13=   0.000 B23=   0.000
REMARK B-factor correction applied to coordinate array B=     1.006
REMARK Bulk solvent: density level= 0.277596 e/A^3, B-factor= 29.219 A^2
REMARK reflections with |Fobs|/sigma_F < 0.0 rejected
REMARK reflections with |Fobs| > 10000 * rms(Fobs) rejected
REMARK theoretical total number of refl. in resol. range:   283454 ( 100.0 % )
REMARK number of unobserved reflections (no entry or 'F'=0):   45349 (  16.0 % )
REMARK number of reflections rejected:                             0 (   0.0 % )
REMARK total number of reflections used:                      238105 (  84.0 % )
REMARK number of reflections in working set:                  214294 (  75.6 % )
REMARK number of reflections in test set:                      23811 (   8.4 % )
CRYST1  401.375  401.375  175.887  90.00  90.00  90.00 P 61 31 2
REMARK VERSION:1.0
ATOM      1  CB  ALA     1     216.271 118.799  27.006  1.00 35.96      8316
ATOM      2  C   ALA     1     216.326 116.084  26.343  1.00 72.05      8316
ATOM      3  O   ALA     1     215.024 117.234  29.489  1.00 72.05      8316
ATOM      4  N   ALA     1     217.536 116.041  27.918  1.00 72.05      8316
ATOM      5  CA  ALA     1     217.068 117.682  28.283  1.00 72.05      8316
ATOM      6  N   ARG     2     216.879 116.080  25.062  1.00111.86      8317
ATOM      7  CA  ARG     2     216.253 114.765  24.160  1.00111.86      8317
ATOM      8  CB  ARG     2     216.718 113.883  23.015  1.00 84.22      8317
ATOM      9  CG  ARG     2     216.965 112.480  23.508  1.00 84.22      8317
ATOM     10  CD  ARG     2     217.603 112.671  24.869  1.00 84.22      8317
ATOM     11  NE  ARG     2     218.902 111.762  24.828  1.00 84.22      8317
ATOM     12  CZ  ARG     2     219.817 111.162  25.630  1.00 84.22      8317
ATOM     13  NH1 ARG     2     219.577 111.177  26.291  1.00 84.22      8317
ATOM     14  NH2 ARG     2     220.180 110.266  30.341  1.00 84.22      8317
ATOM     15  C   ARG     2     216.717 116.819  22.425  1.00111.86      8317
ATOM     16  O   ARG     2     717.888 116.413  22.531  1.00111.86      8317
ATOM     17  N   LYS     3     216.611 116.179  23.388  1.00 99.41      8318
ATOM     18  CA  LYS     3     216.174 116.806  24.651  1.00 99.41      8318
ATOM     19  CB  LYS     3     215.156 115.955  25.748  1.00121.38      8318
ATOM     20  CG  LYS     3     213.728 117.283  25.809  1.00121.38      8318
ATOM     21  CD  LYS     3     213.107 117.031  27.058  1.00121.38      8318
ATOM     22  CE  LYS     3     212.063 118.123  27.365  1.00121.38      8318
ATOM     23  NZ  LYS     3     211.291 117.821  28.594  1.00121.38      8318
ATOM     24  C   LYS     3     217.562 116.275  25.045  1.00 99.41      8318
ATOM     25  O   LYS     3     218.501 117.033  25.281  1.00 99.41      8318
ATOM     26  N   ALA     4     217.622 114.955  25.162  1.00 68.74      8319
ATOM     27  CA  ALA     4     218.849 114.276  25.547  1.00 68.74      8319
ATOM     28  CB  ALA     4     218.721 112.787  25.248  1.00 63.80      8319
ATOM     29  C   ALA     4     220.098 114.837  24.871  1.00 68.74      8319
ATOM     30  O   ALA     4     221.136 115.004  25.518  1.00 68.74      8319
ATOM     31  N   LEU     5     219.997 115.123  23.575  1.00 99.76      8320
ATOM     32  CA  LEU     5     221.137 115.649  22.827  1.00 99.76      8320
ATOM     33  CB  LEU     5     221.074 115.221  21.364  1.00 80.76      8320
ATOM     34  CG  LEU     5     221.106 113.726  21.010  1.00 80.76      8320
ATOM     35  CD1 LEU     5     221.682 113.588  19.548  1.00 80.76      8320
ATOM     36  CD2 LEU     5     222.114 112.971  21.988  1.00 80.76      8320
ATOM     37  C   LEU     5     221.138 117.162  22.923  1.00 80.76      8320
ATOM     38  O   LEU     5     221.954 117.744  22.098  1.00 80.76      8320
ATOM     39  N   ILE     6     220.689 117.757  23.916  1.00 53.20      8321
ATOM     40  CA  ILE     6     220.688 119.181  24.120  1.00 53.20      8321
ATOM     41  CB  ILE     6     219.574 119.775  24.975  1.00 53.33      8321
ATOM     42  CG2 ILE     6     219 887 121.471  24.971  1.00 53.33      8321
ATOM     43  CG1 ILE     6     218.748 119.648  23.923  1.00 53.33      8321
ATOM     44  CD1 ILE     6     217.589 120.564  23.707  1.00 53.33      8321
ATOM     45  C   ILE     6     222.032 119.233  24.644  1.00 82.20      8321
ATOM     46  O   ILE     6     221.177 119.795  25.467  1.00 82.20      8321
ATOM     47  N   GLU     7     222.760 118.557  25.425  1.00 93.58      8322
ATOM     48  CA  GLU     7     223.736 118.479  26.510  1.00 93.58      8322
ATOM     49  CB  GLU     7     224.758 116.998  26.589  1.00 92.66      8322
ATOM     50  CG  GLU     7     225.213 116.621  25.452  2.00 92.66      8322
ATOM     51  CD  GLU     7     225.278 115.110  25.168  1.00 92.66      8322
ATOM     52  OE1 GLU     7     225.547 114.325  26.092  1.00 92.66      8322
ATOM     53  OE2 GLU     7     224.865 114.727  24.045  1.00 92.66      8322
ATOM     54  C   GLU     7     224.849 119.341  26.027  1.00 93.58      8322
ATOM     55  O   GLU     7     225.992 119.249  26.467  1.00 93.58      8322
ATOM     56  N   LYS     8     224.492 120.194  25.070  1.00 95.44      8323
ATOM     57  CA  LYS     8     223.611 121.339  24.492  1.00 95.44      8323
ATOM     58  CB  LYS     8     224.824 121.732  23.281  1.00 92.82      8323
ATOM     59  CG  LYS     8     225.681 122.802  22.523  1.00 92.82      8323
ATOM     60  CD  LYS     8     225.536 124.187  23.152  1.00 92.82      8324
ATOM     61  CE  LYS     8     226.457 125.196  22.502  1.00 92.82      8324
ATOM     62  NZ  LYS     8     226 194 126.553  22.942  1.00 92.82      8324
ATOM     63  C   LYS     8     225.579 122.768  25.529  1.00 95.44      8324
ATOM     64  O   LYS     8     226.602 123.950  25.571  1.00 95.44      8324
ATOM     65  N   ALA     9     224.586 123.415  26.352  1.00 59.01      8325
ATOM     66  CA  ALA     9     224.583 123.016  27.596  1.00 59.01      8325
ATOM     67  CB  ALA     9     223.168 123.699  28.078  1.00 36.06      8325
ATOM     68  C   ALA     9     225.651 122.931  28.643  1.00 59.01      8325
ATOM     69  O   ALA     9     225.991 123.899  29.303  1.00 59.01      8325
ATOM     70  N   LYS    10     226.589 121.599  28.588  1.00128.35      8326
ATOM     71  CA  LYS    10     226.526 120.981  29.627  1.00128.35      8326
ATOM     72  CB  LYS    10     226.961 119.481  29.473  1.00 72.22      8326
ATOM     73  CG  LYS    10     225.696 118.696  29.660  1.00 72.22      8326
ATOM     74  CD  LYS    10     225.696 117.195  29.995  1.00 72.22      8326
ATOM     75  CE  LYS    10     224.506 116.336  30.392  1.00 72.22      8326
ATOM     76  NZ  LYS    10     224.820 114.888  40.248  1.00 72.22      8326
ATOM     77  C   LYS    10     227.948 121.474  30.618  1.00128.35      8326
ATOM     78  O   LYS    10     228.979 120.783  30.681  1.00128.35      8326
ATOM     79  N   ARG    11     228.179 122.675  29.085  1.00129.42      8327
ATOM     80  CA  ARG    11     229.509 123.242  29.027  1.00129.42      8327
ATOM     81  CB  ARG    11     230.084 123.403  30.468  1.00 99.37      8328
ATOM     82  CG  ARG    11     229.765 124.774  31.053  1.00 99.37      8328
ATOM     83  CD  ARG    11     230.231 125.813  30.044  1.00 99.37      8328
ATOM     84  NE  ARG    11     229.457 127.044  30.136  1.00 99.37      8328
ATOM     85  CZ  ARG    11     229.581 128.049  30.982  1.00 99.37      8329
ATOM     86  NH1 ARG    11     230.387 127.976  32.037  1.00 99.37      8329
ATOM     87  NH2 ARG    11     228.888 129.137  30.637  1.00 99.37      8329
ATOM     88  C   ARG    11     230.477 122.456  28.166  1.00133.42      8329
ATOM     89  O   ARG    11     230.268 123.246  26.976  1.00133.42      8329
ATOM     90  N   TYR    12     231.545 122.913  28.784  1.00128.36      8330
ATOM     91  CA  TYR    12     232.606 123.716  28.079  1.00128.36      8330
ATOM     92  CB  TYR    12     233.963 123.991  28.781  1.00128.36      8330
ATOM     93  CG1 TYR    12     233.984 123.716  27.348  1.00128.36      8330
ATOM     94  CG2 TYR    12     235.103 124.102  27.002  1.00128.36      8330
ATOM     95  C   TYR    12     232.162 125.144  28.331  1.00125.62      8330
ATOM     96  O   TYR    12     233.352 119.574  29.558  1.00105.62      8330
ATOM     97  N   PRO    13     232.225 118.899  27.728  1.00201.09      8331
ATOM     98  CA  PRO    13     231.938 119.047  26.650  1.00132.01      8331
ATOM     99  CB  PRO    13     232.370 117.696  26.113  1.00132.01      8331
ATOM    100  CG  PRO    13     231.575 116.751  26.725  1.00132.01      8331
ATOM    101  C   PRO    13     233.057 117.118  26.100  1.00201.09      8331
ATOM    102  O   PRO    13     234 324 116.418  27.298  1.00201.09      8331
ATOM    103  N   LYS    14     234.577 117.624  25.099  1.00115.47      8332
ATOM    104  CA  LYS    14     236 012 117.418  25 282  1.00115.47      8332
ATOM    105  CB  LYS    14     234.462 136 001  38.870  1.00313.29      8316
ATOM    106  CG  LYS    14     235.976 136.836  39.772  1.00313.29      8316
ATOM    107  CD  LYS    14     236.389 136.365  39.388  1.00313.29      8316
ATOM    108  CE  LYS    14     236.255 135.829  48.071  1.00313.29      8316
ATOM    109  NZ  LYS    14     235.159 131.914  39.629  1.00313.29      8316
ATOM    110  C   LYS    14     236.874 138.439  38.513  1.00115.41      8316
ATOM    111  O   LYS    14     237.978 139.278  39.158  1.00115.41      8316
ATOM    112  N   PHE    15     236.811 138.436  37.404  1.00 59.11      8316
ATOM    113  CA  PHE    15     237.633 139.342  36.592  1.00 59.11      8316
ATOM    114  CB  PHE    15     238.528 138.523  35.659  1.00 66.69      8316
ATOM    115  CG  PHE    15     238.762 137.110  36.936  1.00 66.69      8316
ATOM    116  CD1 PHE    15     237.757 136.167  35.896  1.00 66.69      8316
ATOM    117  CD2 PHE    15     239.970 136.747  36.520  1.00 66.69      8316
ATOM    118  CE1 PHE    15     237.950 134.845  36.293  1.00 66.69      8316
ATOM    119  CE2 PHE    15     240.174 135.445  36.991  1.00 66.69      8316
ATOM    120  CZ  PHE    15     239.159 134.496  36.874  1.00 66.69      8316
ATOM    121  N   PHE    15     236.767 139.397  35.496  1.00 59.11      8316
ATOM    122  C   PHE    15     235.693 139.921  35.156  1.00 59.11      8316
ATOM    123  O   PHE    15     237.829 141.525  36.999  1.00100.69      8316
ATOM    124  N   LYS    16     236.427 142.523  34.713  1.00100.69      8316
ATOM    125  CA  LYS    16     236.882 123.940  35.125  1.00 98.21      8316
ATOM    126  CB  LYS    16     238.167 124.432  34.488  1.00 98.21      8316
ATOM    127  CG  LYS    16     237.931 125.798  33.762  1.00 98.21      8316
ATOM    128  CD  LYS    16     237.266 126.782  34.686  1.00 98.21      8316
ATOM    129  CE  LYS    16     237.044 128.097  33.987  1.00 98.21      8316
ATOM    130  NZ  LYS    16     236.291 129.475  33.195  1.00100.69      8316
ATOM    131  C   LYS    16     235.840 123.448  32.578  1.00100.69      8316
ATOM    132  O   LYS    16     236.627 121.349  32.586  1.00121.79      8316
ATOM    133  N   VAL    17     238.189 121.228  31.229  1.00121.79      8316
ATOM    134  CA  VAL    17     237.658 120.456  30.506  1.00 79.35      8316
ATOM    135  CB  VAL    17     238.890 121.428  30.232  1.00 79.35      8316
ATOM    136  CG2 VAL    17     238.112 119.116  31.435  1.00 79.35      8316
ATOM    137  CG1 VAL    17     235.147 120.548  30.787  1.00121.79      8316
ATOM    138  C   VAL    17     234.881 120.314  29.613  1.00121.79      8316
ATOM    139  O   VAL    17     234.388 120.228  31.818  1.00 72.30      8316
ATOM    140  N   ARG    18     233.100 119.572  31.642  1.00 72.30      8316
ATOM    141  CA  ARG    18     232.999 118.379  30.599  1.00 59.34      8316
ATOM    142  CB  ARG    18     234.199 117.432  32.394  1.00 59.34      8316
ATOM    143  CG  ARG    18     233.940 116.257  33.342  1.00 59.34      8316
ATOM    144  CD  ARG    18     234.887 115.162  33.337  1.00 59.34      8316
ATOM    145  NE  ARG    18     234.910 114.032  34.046  1.00 59.34      8316
ATOM    146  CZ  ARG    18     834.634 113.813  35.023  1.00 59.34      8316
ATOM    147  NH1 ARG    18     835.808 113.094  33.779  1.00 59.34      8316
ATOM    148  NH2 ARG    18     232.939 128.537  31.931  1.00 72.30      8316
ATOM    149  C   ARG    18     231.183 128.350  32.858  1.00 72.30      8316
ATOM    150  O   ARG    18     231.767 121.993  31.058  1.00123.42      8316
ATOM    151  N   ALA    19     230.663 122.538  31.318  1.00123.42      8316
ATOM    152  CA  ALA    19     231.272 123.868  31.650  1.00 34.95      8316
ATOM    153  CB  ALA    19     229.973 123.685  29.915  1.00123.42      8316
ATOM    154  C   ALA    19     230.568 122.773  30.865  1.00123.42      8316
ATOM    155  O   ALA    19     729.590 122.807  29.932  1.00 92.41      8316
ATOM    156  N   THR    20     227.914 122.951  29.660  1.00 92.41      8316
ATOM    157  CA  THR    20     227.301 121.627  29.189  1.00 64.67      8316
ATOM    158  CB  THR    20     226.867 120.823  30.318  1.00 64.67      8316
ATOM    159  CG2 THR    20     228.982 119.956  29.051  1.00 64.67      8316
ATOM    160  OG1 THR    20     227.788 119.956  30.871  1.00 64.67      8316
ATOM    161  C   THR    20     227.415 118.898  30.815  1.00 64.67      8316
ATOM    162  O   THR    20     228.932 120.229  29.288  1.00 64.67      8316
ATOM    163  N   THR    21     225.125 119.188  30.994  1.00 64.67      8316
ATOM    164  CA  THR    21     226.077 118.502  30.953  1.00 64.67      8316
ATOM    165  CB  THR    21     225.691 117.411  31.613  1.00 64.67      8316
ATOM    166  C   THR    21     226.940 120.061  29.228  1.00 92.41      8316
ATOM    167  O   THR    21     226.397 121.340  27.185  1.00 92.41      8316
ATOM    168  N   THR    21     226.874 125.126  29.064  1.00103.54      8316
ATOM    169  CA  THR    21     225.977 126.248  28.977  1.00103.54      8316
ATOM    170  CB  THR    21     226.962 126.869  27.275  1.00 70.67      8316
ATOM    171  OG1 THR    21     227.984 126.718  27.997  1.00 70.67      8316
ATOM    172  CG2 THR    21     226.816 128.329  27.145  1.00 70.67      8316
ATOM    173  C   THR    21     224.472 126.685  28.334  1.00103.54      8316
ATOM    174  O   THR    21     223.887 126.697  27.782  1.00103.54      8316
ATOM    175  N   VAL    22     224.974 127.844  28.706  1.00 67.95      8316
ATOM    176  CA  VAL    22     222.501 126.552  28.519  1.00 67.95      8316
ATOM    177  CB  VAL    22     222.000 123.784  29.679  1.00 65.88      8316
ATOM    178  CG1 VAL    22     223.694 122.384  29.859  1.00 65.88      8316
ATOM    179  CG2 VAL    22     221.904 121.303  29.176  1.00 65.88      8316
ATOM    180  C   VAL    22     222.399 123.987  29.844  1.00 65.88      8316
ATOM    181  O   VAL    22     221.970 110.862  29.988  1.00 65.88      8316
ATOM    182  N   ARG    23     221.040 110.918  29.057  1.00 67.95      8316
ATOM    183  CA  ARG    23     222.466 112.649  29.373  1.00 67.95      8316
ATOM    184  CB  ARG    23     221.501 125.792  28.488  1.00 67.95      8316
ATOM    185  CG  ARG    23     221.879 126.806  29.069  1.00 67.95      8316
ATOM    186  CD  ARG    23     220.545 125.795  27.582  1.00 96.33      8316
ATOM    187  NE  ARG    23     219.612 125.823  27.412  1.00 96.33      8316
ATOM    188  CZ  ARG    23     218.486 126.480  26.619  1.00 55.35      8316
ATOM    189  NH1 ARG    23     217.318 127.048  25.957  1.00 55.35      8316
ATOM    190  NH2 ARG    23     219.016 127.093  28.774  1.00 96.33      8316
ATOM    191  C   ARG    23     218.427 126.165  29.487  1.00 67.26      8316
ATOM    192  N   VAL    24     218.958 128.263  29.162  1.00 98.57      8316
ATOM    193  CA  VAL    24     218.487 129.722  29.037  1.00 98.57      8316
ATOM    194  CB  VAL    24     218.427 130.246  30.637  1.00 64.18      8316
ATOM    195  CG1 VAL    24     217.100 130.729  31.232  1.00 64.18      8316
ATOM    196  CG2 VAL    24     219.588 130.617  31.357  1.00 64.18      8316
ATOM    197  C   VAL    24     216.987 129.299  30.388  1.00 98.57      8316
ATOM    198  O   VAL    24     216.621 127.735  31.681  1.00 98.57      8316
ATOM    199  N   ARG    25     216.186 129.278  29.327  1.00 75.21      8316
ATOM    200  CA  ARG    25     213.849 128.938  29.240  1.00 67.26      8316
ATOM    201  CB  ARG    25     212.528 129.096  27.786  1.00 67.26      8316
ATOM    202  CG  ARG    25     212.570 130.230  27.835  1.00 67.26      8316
ATOM    203  CD  ARG    25     211.948 131.540  28.324  1.00 67.26      8316
ATOM    204  NE  ARG    25     211.011 131.184  26.007  1.00 67.26      8316
ATOM    205  CZ  ARG    25     210.602 132.069  26.911  1.00 67.26      8316
ATOM    206  NH1 ARG    25     210.485 131.193  24.795  1.00 67.26      8316
ATOM    207  NH2 ARG    25     214.482 128.524  29.578  1.00 75.21      8316
ATOM    208  C   ARG    25     213.399 129.935  29.373  1.00 75.21      8316
ATOM    209  O   ARG    25     216.349 126.067  27.873  1.00 68.35      8316
ATOM    210  N   GLY    26     214.736 126.799  27.325  1.00 68.35      8316
ATOM    211  CA  GLY    26     214.636 125.959  29.801  1.00 19.43      8316
ATOM    212  C   GLY    26     215.950 125.888  25.005  1.00 19.43      8316
ATOM    213  O   GLY    26     215.799 127.704  27.645  1.00 68.35      8316
ATOM    214  N   GLY    27     213.584 127.519  27.490  1.00 77.36      8316
ATOM    215  CA  GLY    27     214.277 126.311  29.113  1.00 77.36      8316
ATOM    216  C   GLY    27     210.020 123.154  29.450  1.00 77.36      8316
ATOM    217  O   GLY    27     214.723 122.684  27.188  1.00 77.36      8316
ATOM    218  N   GLY    28     219.663 121.097  27.226  1.00 78.05      8316
ATOM    219  CA  GLY    28     218.945 121.316  25.963  1.00 78.05      8316
ATOM    220  C   GLY    28     218.792 121.884  24.744  1.00 78.05      8316
ATOM    221  O   GLY    28     218.084 123.682  27.670  1.00 78.05      8316
ATOM    222  N   ARG    29     218.449 123.258  27.977  1.00 78.05      8316
ATOM    223  CA  ARG    29     217.774 121.946  27.907  1.00 78.05      8316
ATOM    224  CB  ARG    29     218.424 121.293  28.780  1.00 78.05      8316
ATOM    225  CG  ARG    29     218.432 121.843  19.590  1.00 78.05      8316
ATOM    226  CD  ARG    29     219.340 123.057  19.353  1.00 78.05      8316
ATOM    227  NE  ARG    29     218.940 123.687  19.187  1.00 78.05      8316
ATOM    228  CZ  ARG    29     219.239 123.193  18.439  1.00 79.05      8316
ATOM    229  NH1 ARG    29     220.284 123.784  26.722  1.00145.74      8316
ATOM    230  NH2 ARG    29     218.782 124.213  24.209  1.00145.74      8316
ATOM    231  C   ARG    29     229.702 124.213  24.209  1.00145.74      8316
ATOM    232  O   ARG    29     221.082 122.377  25.231  1.00 71.56      8316
ATOM    233  N   ALA    30     227.330 122.437  25.295  1.00 71.56      8316
ATOM    234  CA  ALA    30     226.078 122.826  24.975  1.00 71.56      8316
ATOM    235  CB  ALA    30     227.194 121.353  26.078  1.00 74.10      8316
ATOM    236  C   ALA    30     223.150 123.359  23.899  1.00 71.56      8316
ATOM    237  O   ALA    30     226 197 123.180  23.729  1.00 71.56      8316
ATOM    238  N   ALA    31     227.676 121.987  22.909  1.00 74.10      8316
ATOM    239  CA  ALA    31     227.950 121.581  21.588  1.00 74.10      8316
ATOM    240  CB  ALA    31     231.823 121.369  29.639  1.00 60.45      8316
ATOM    241  C   ALA    31     232.558 120.656  29.389  1.00 60.45      8316
ATOM    242  O   ALA    31     233.954 119.294  29.683  1.00 60.45      8316
ATOM    243  N   ARG    32     231.251 118.755  29.947  1.00 60.45      8316
ATOM    244  CA  ARG    32     229.364 119.908  30.913  1.00 74.10      8316
ATOM    245  CB  ARG    32     229.174 116.780  29.597  1.00 60.45      8316
ATOM    246  CG  ARG    32     225.880 117.696  30.222  1.00 74.10      8316
ATOM    247  CD  ARG    32     225.413 119.010  29.610  1.00 74.10      8316
ATOM    248  NE  ARG    32     224 823 120.110  29.066  1.00 74.10      8316
ATOM    249  CZ  ARG    31     222.729 121.305  29.962  1.00 43.32      8316
ATOM    250  CA  ARG    31     223.182 125.989  30.768  1.00 43.32      8316
```



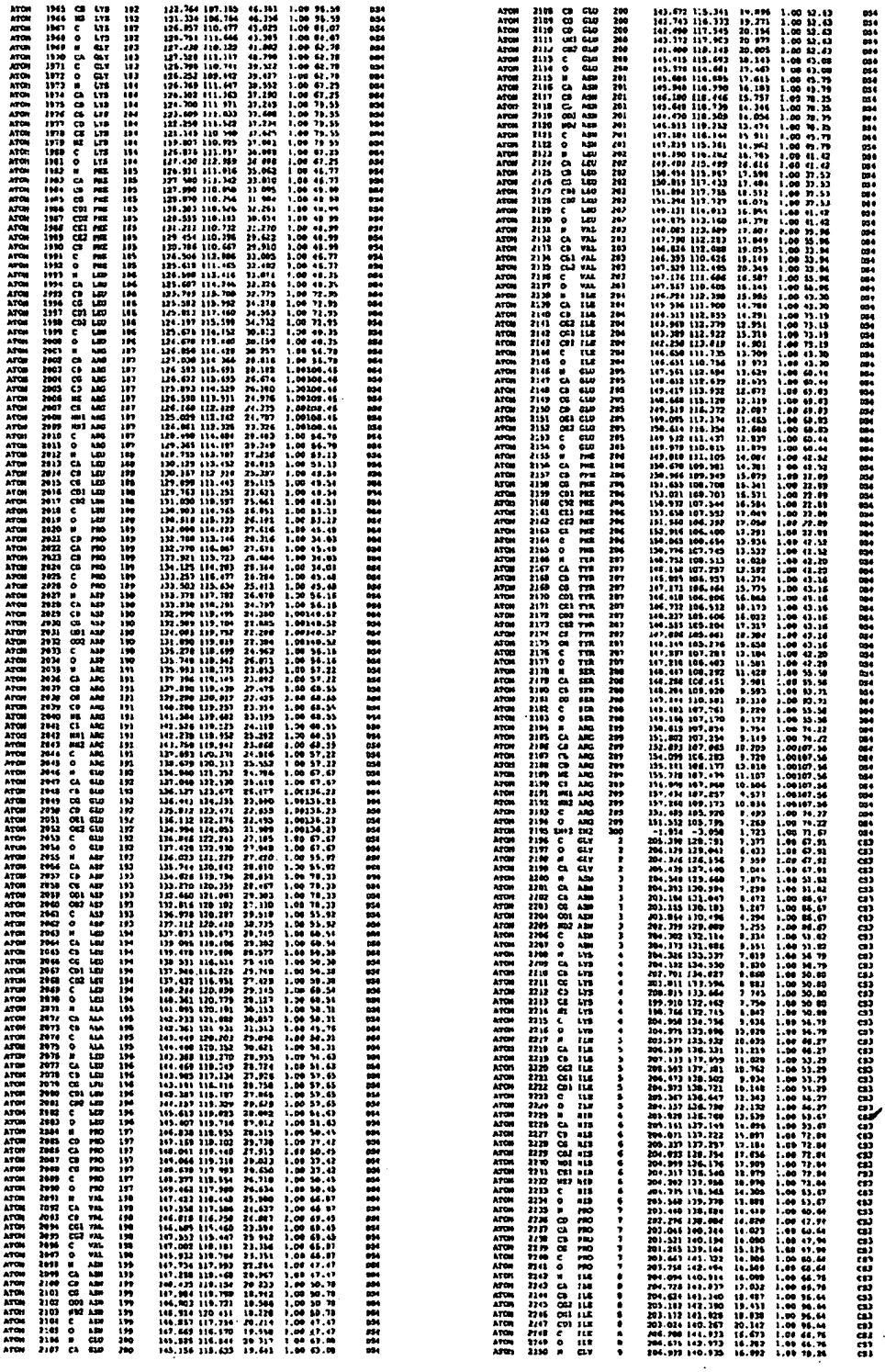


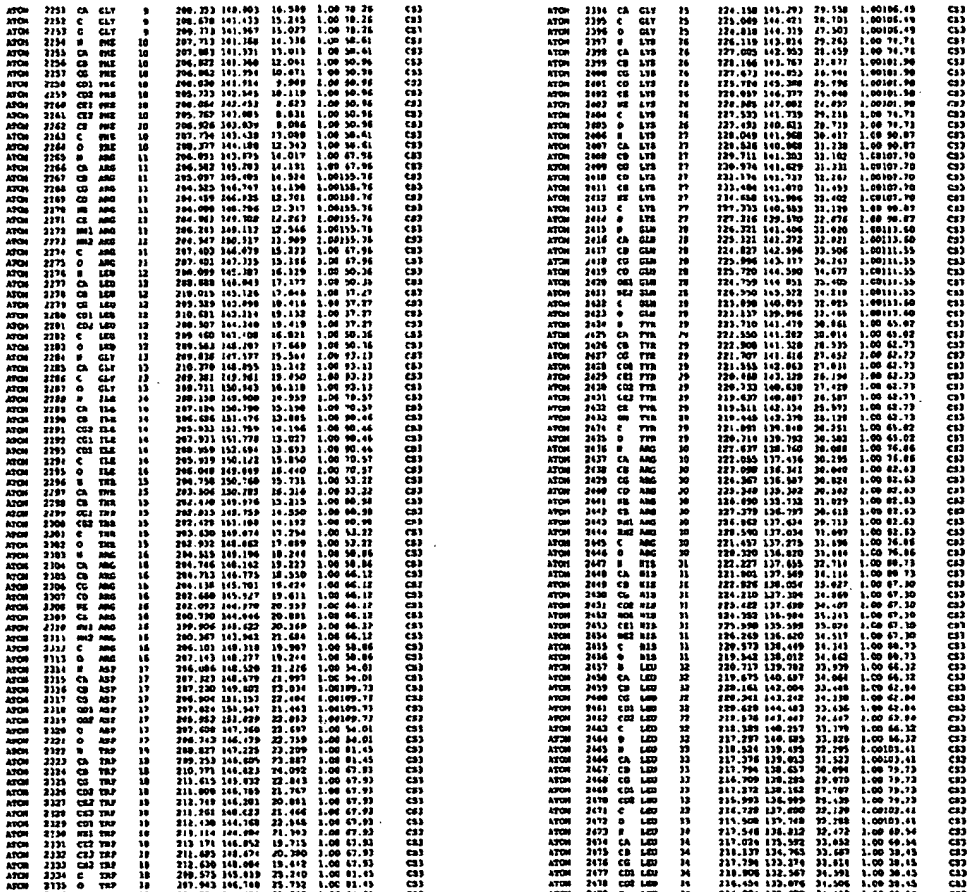


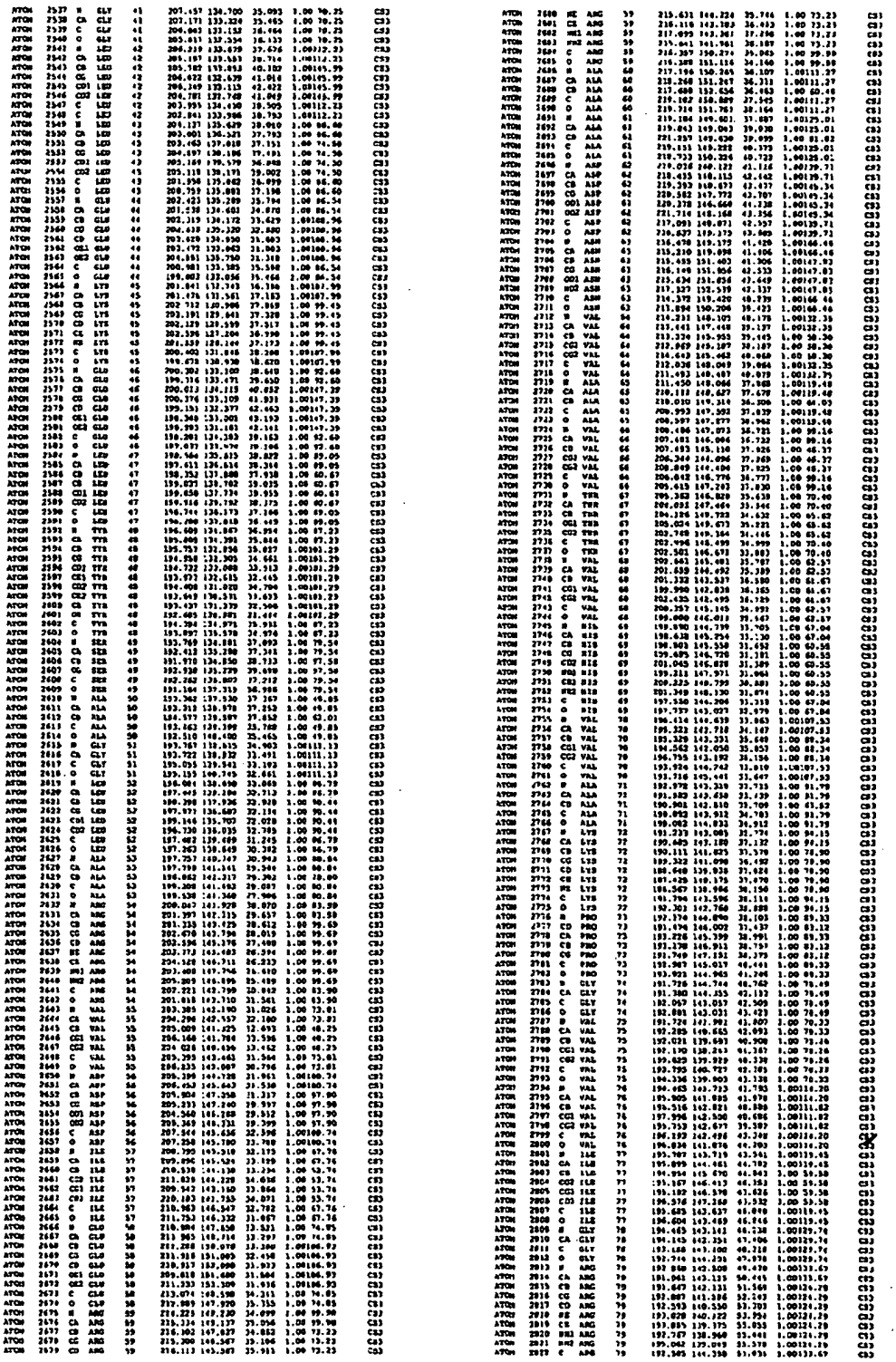


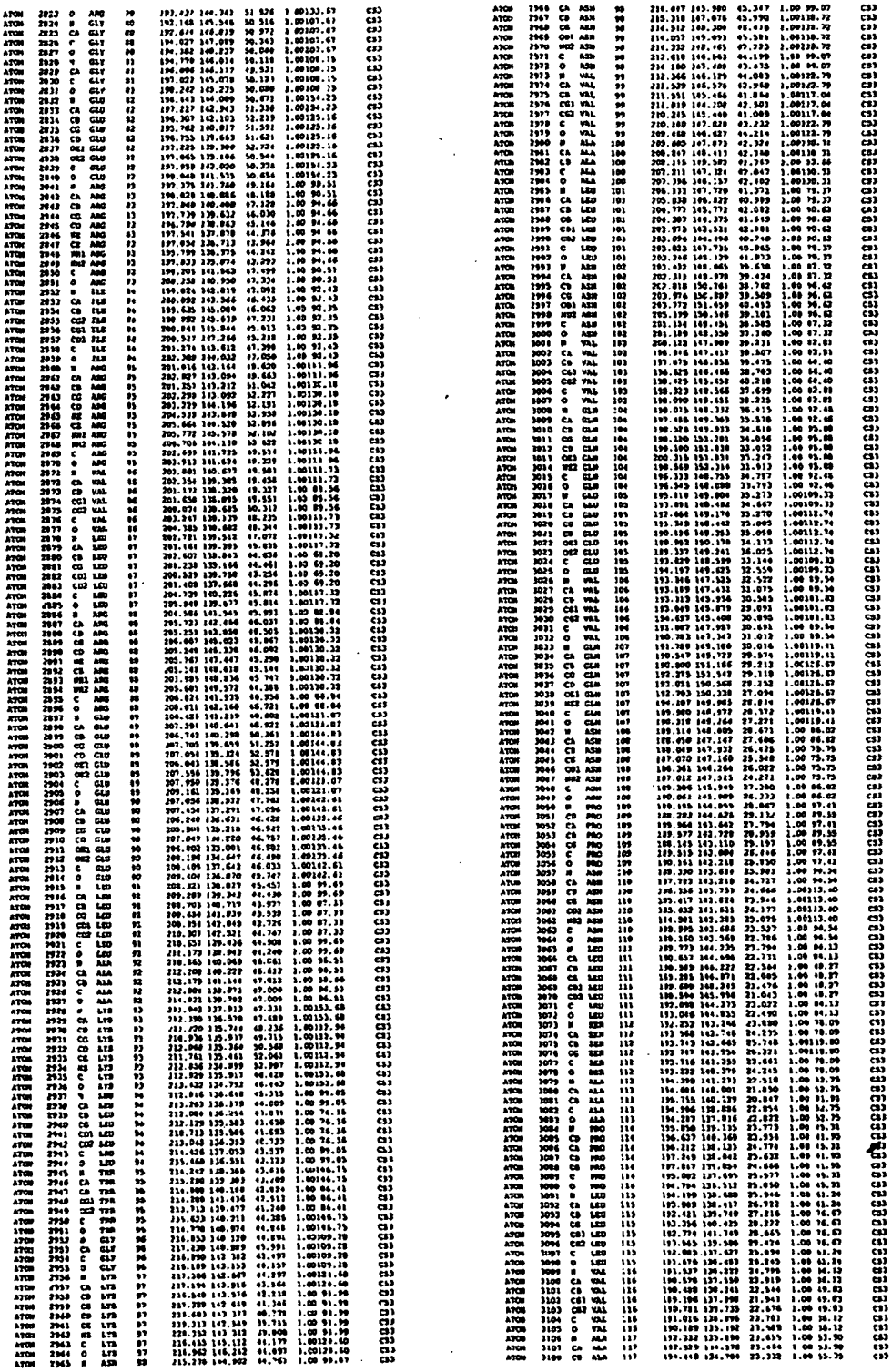


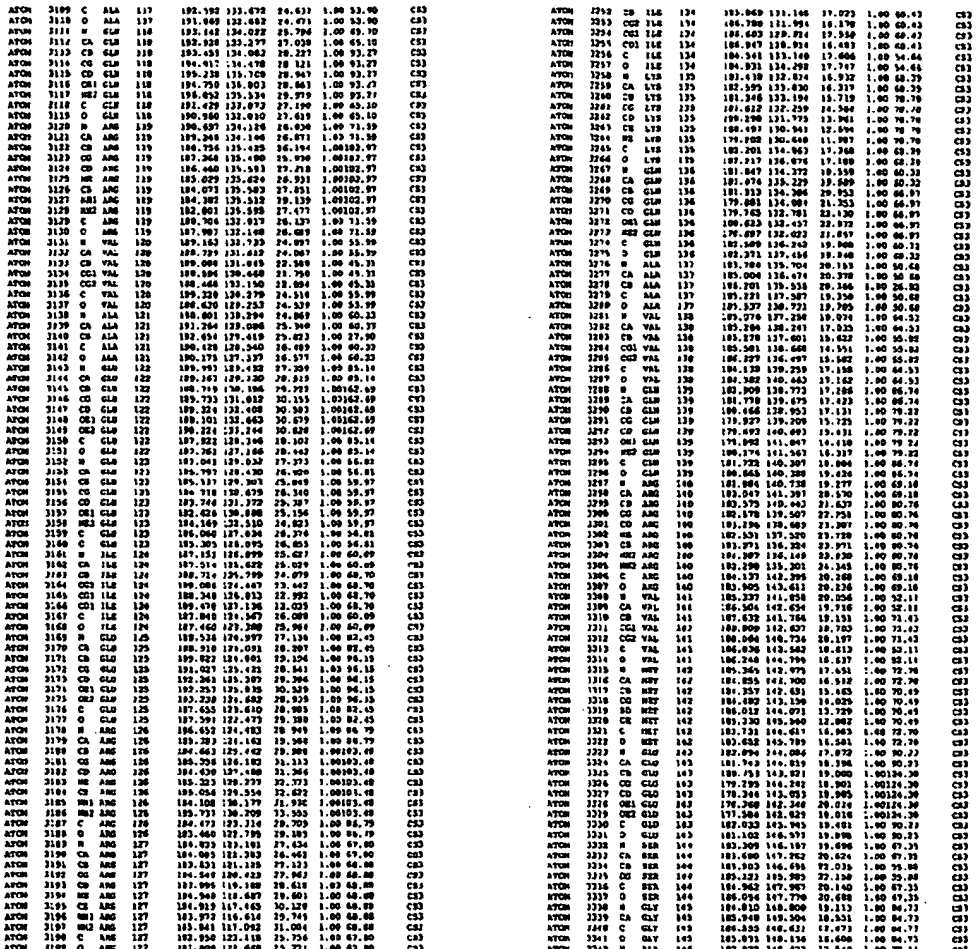


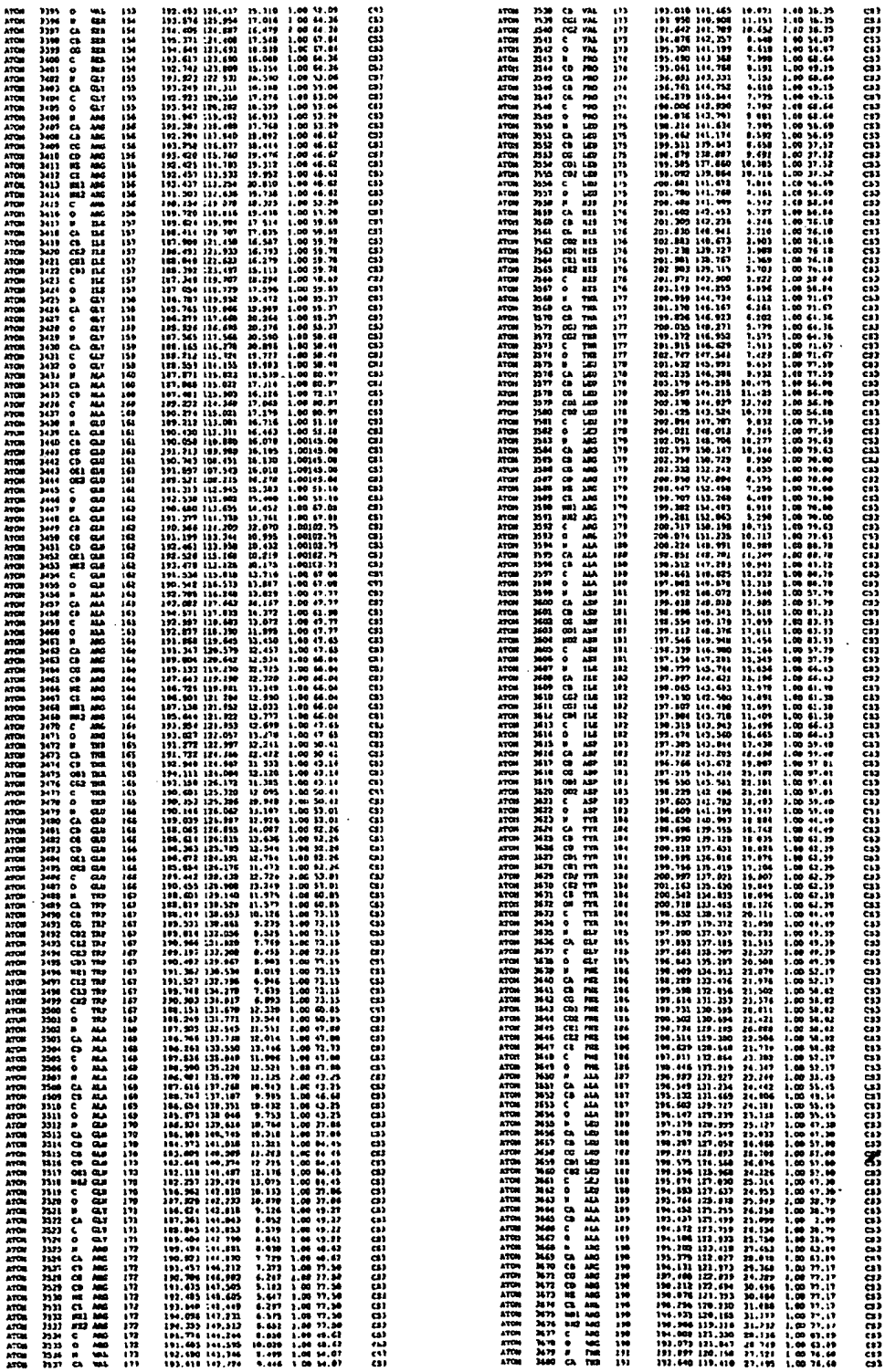


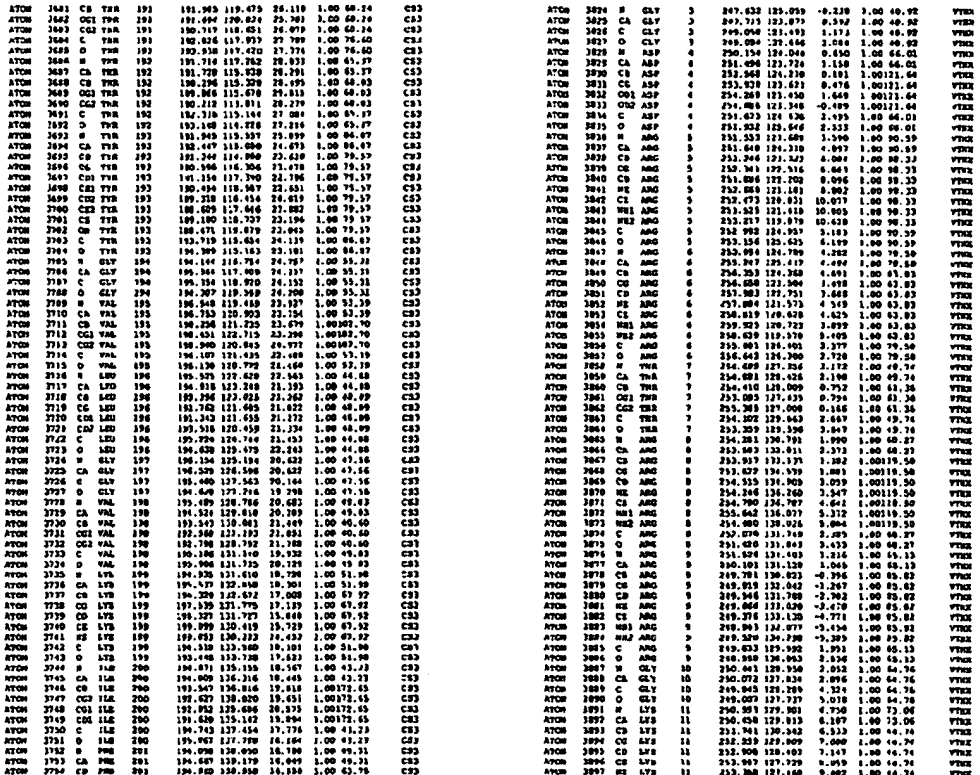


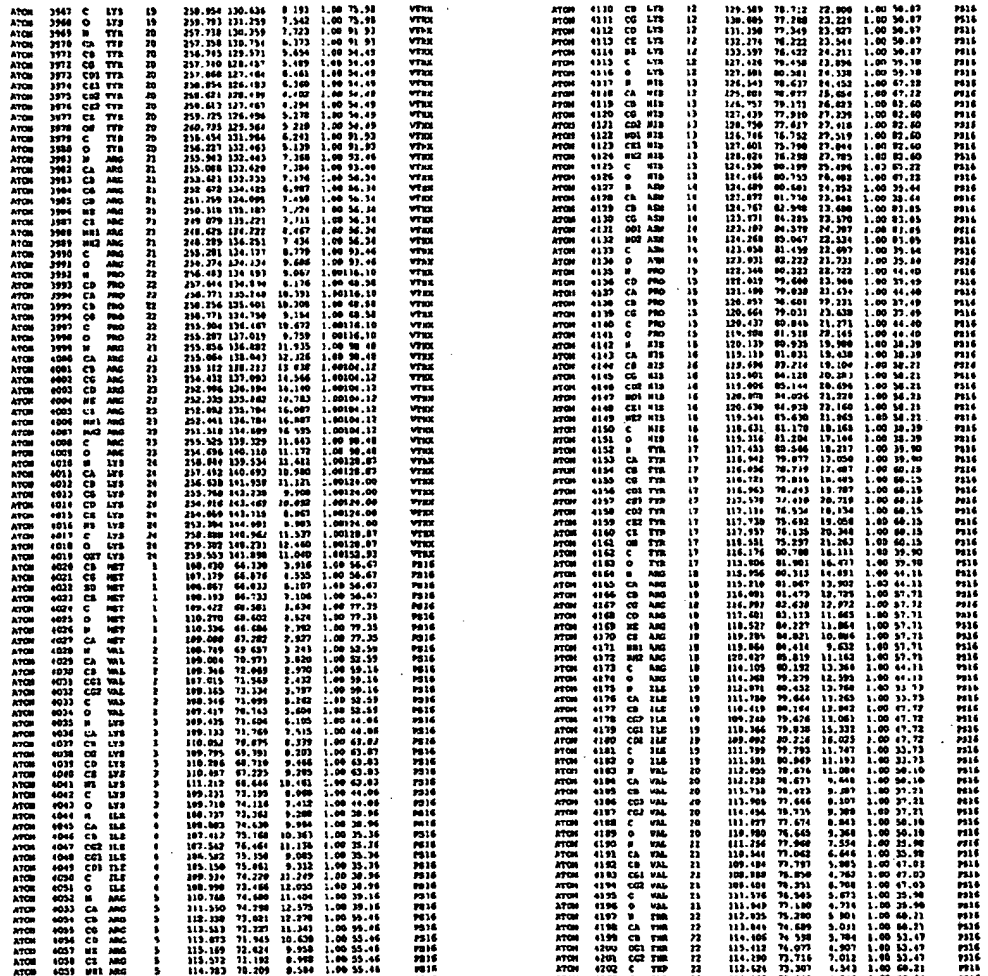


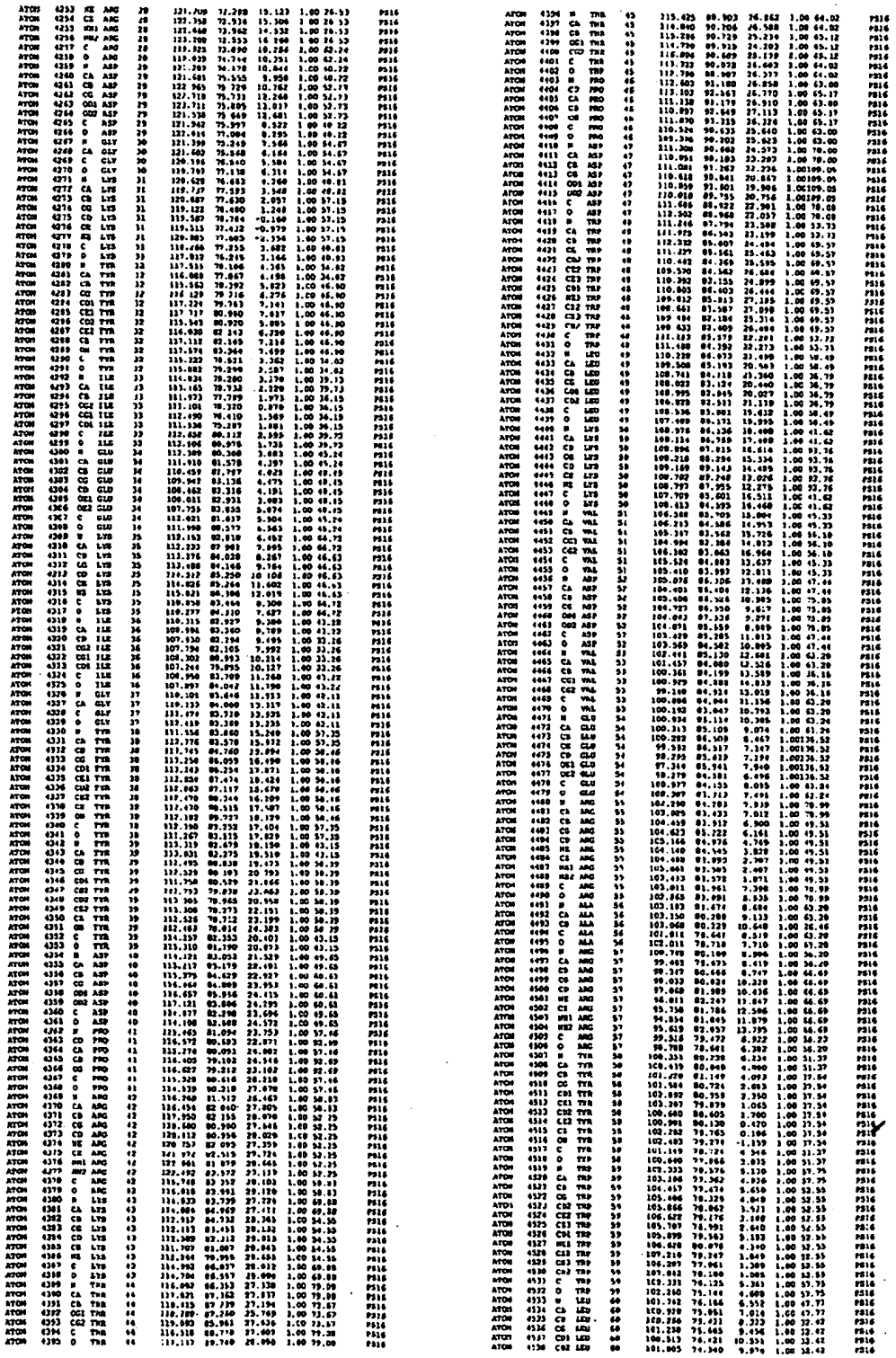




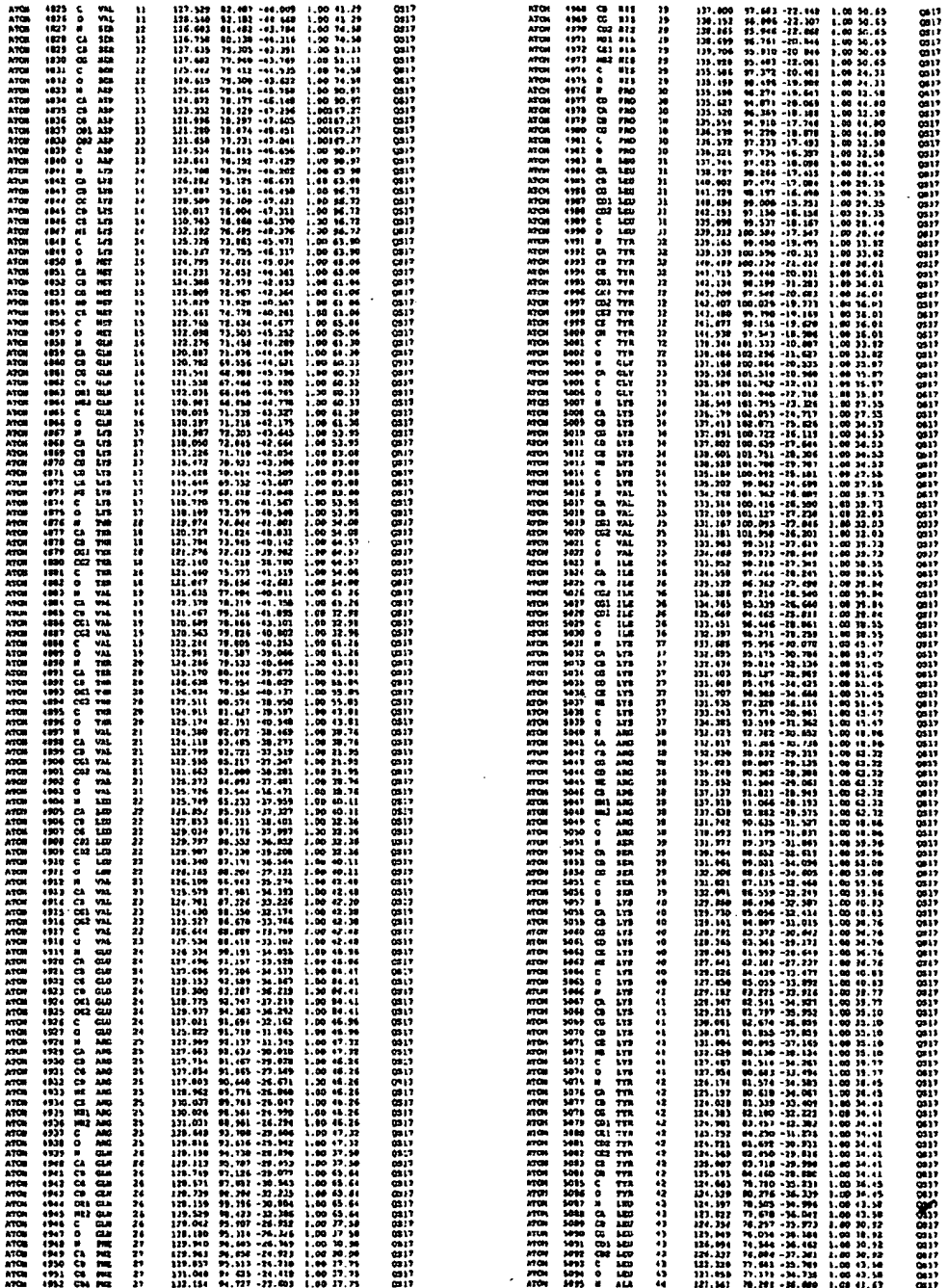




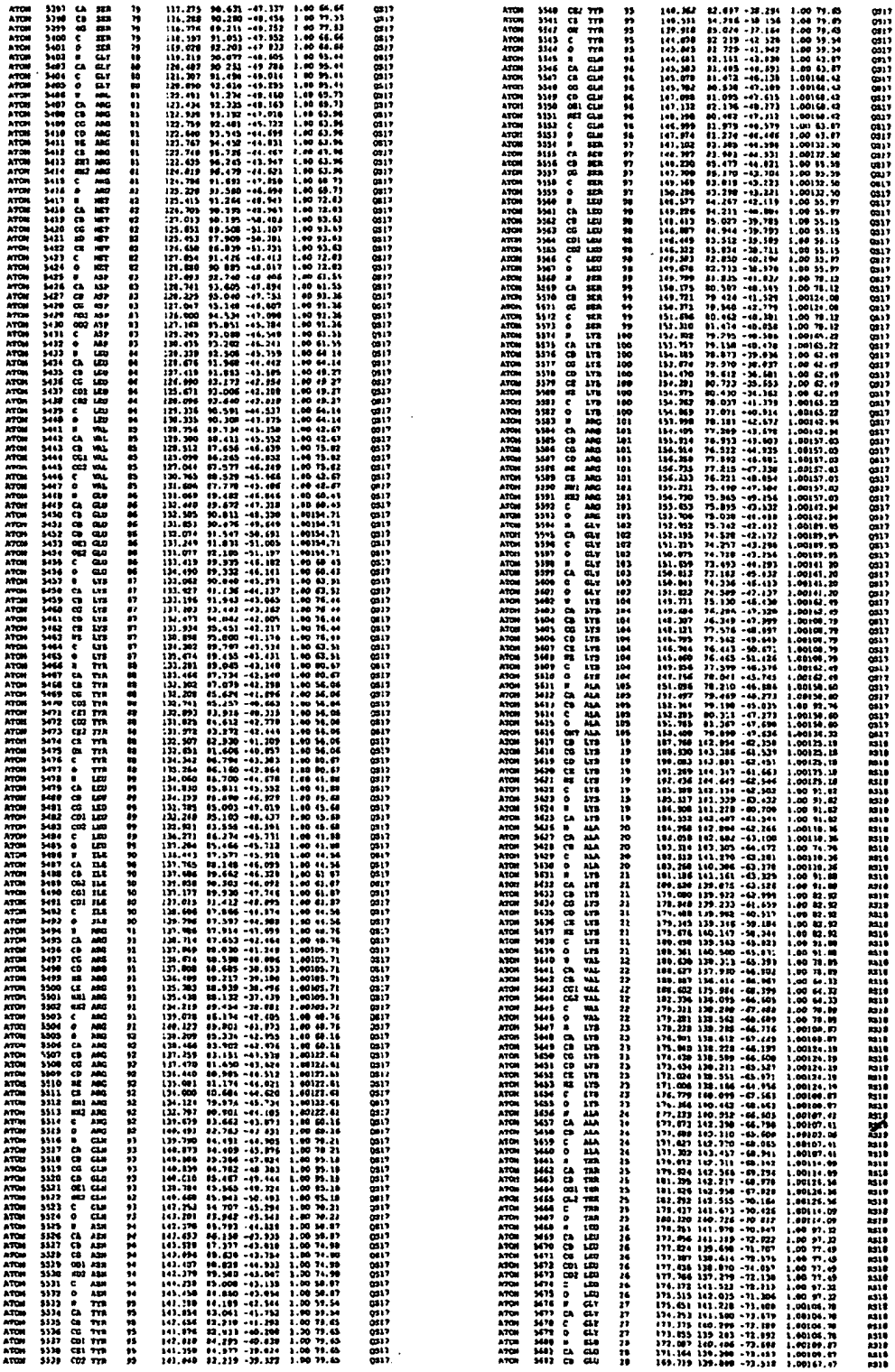


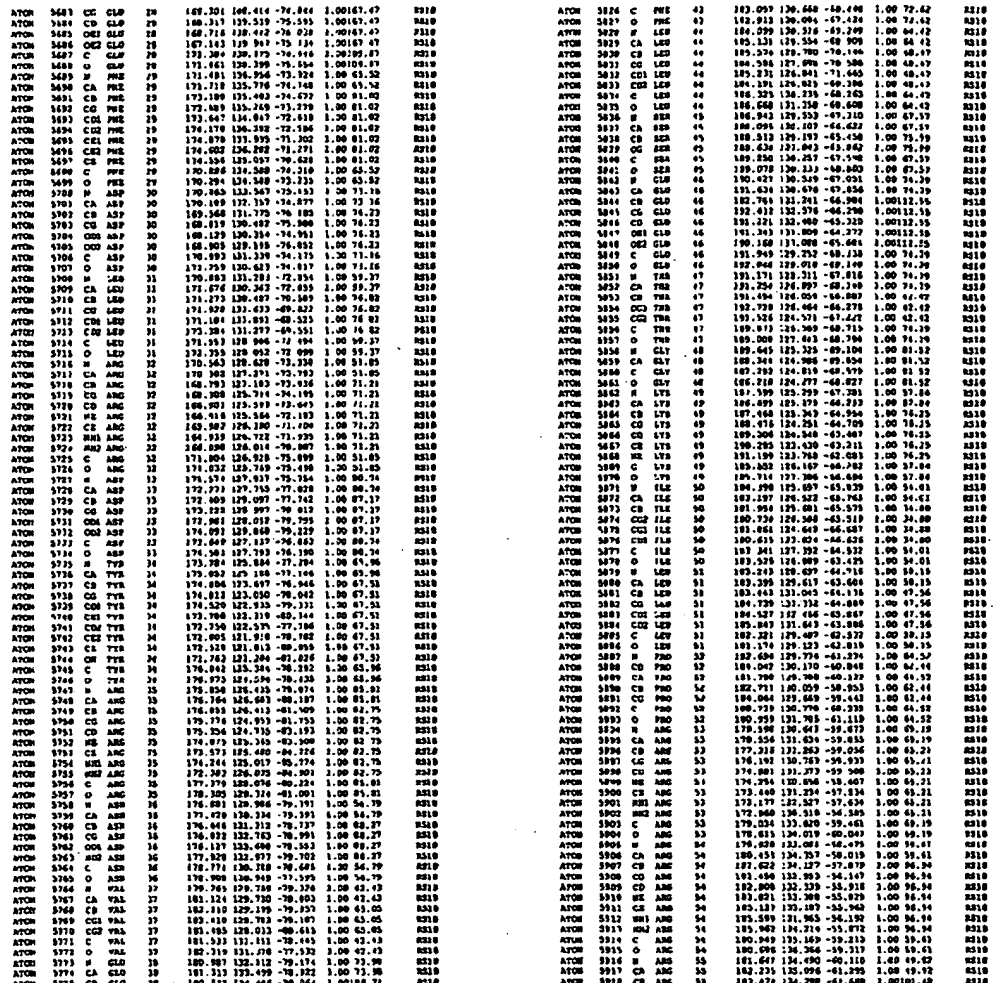


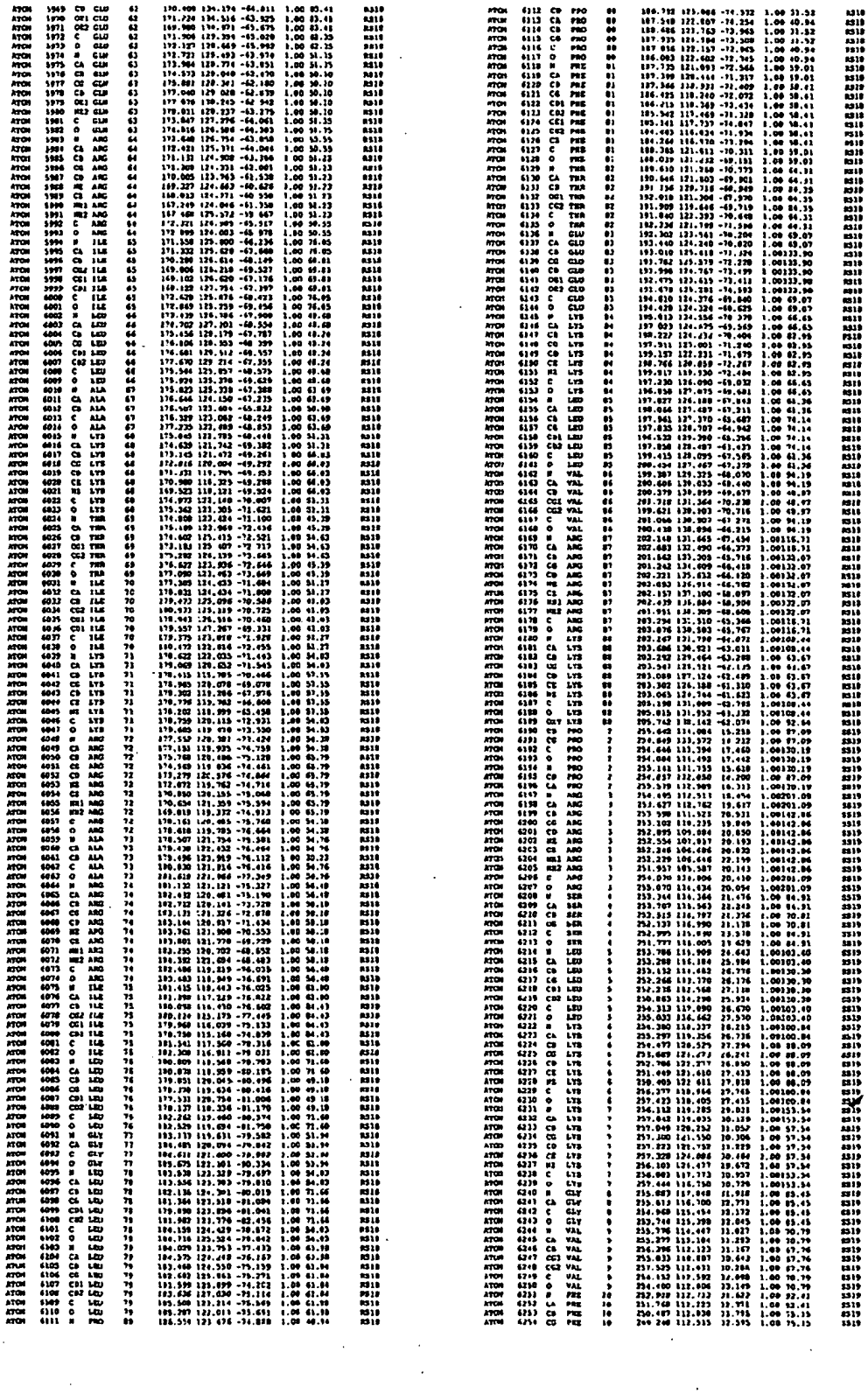


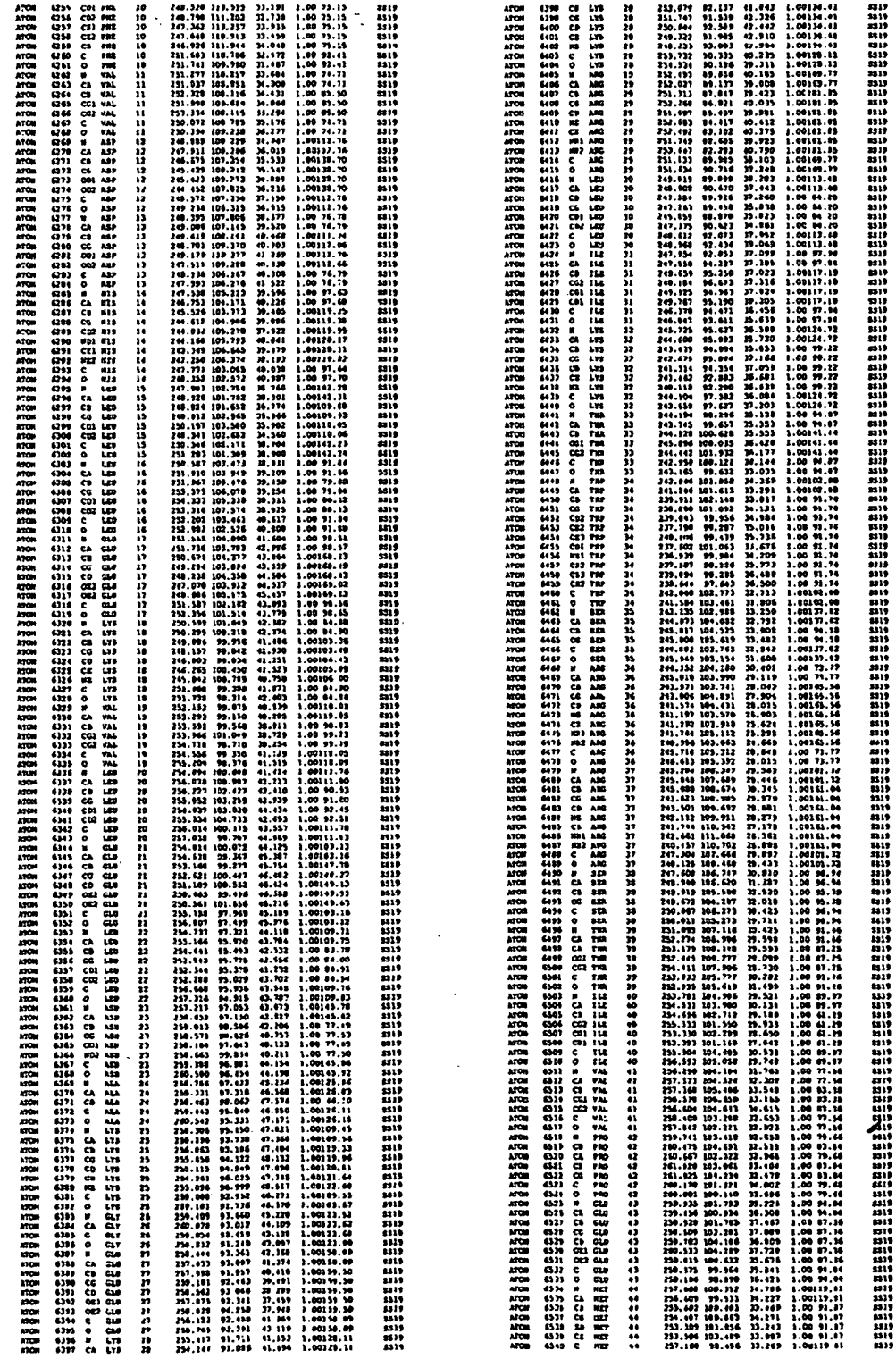


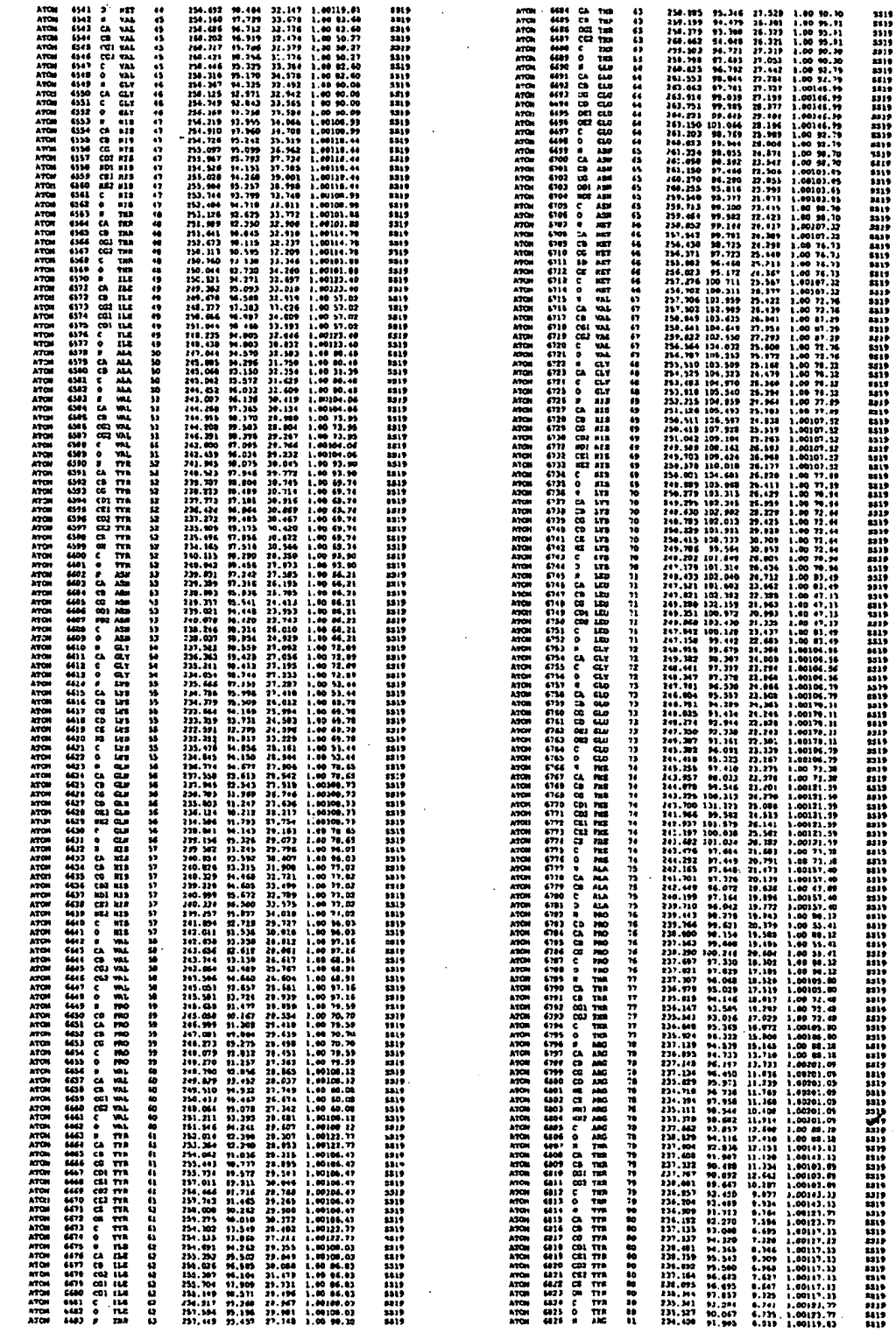


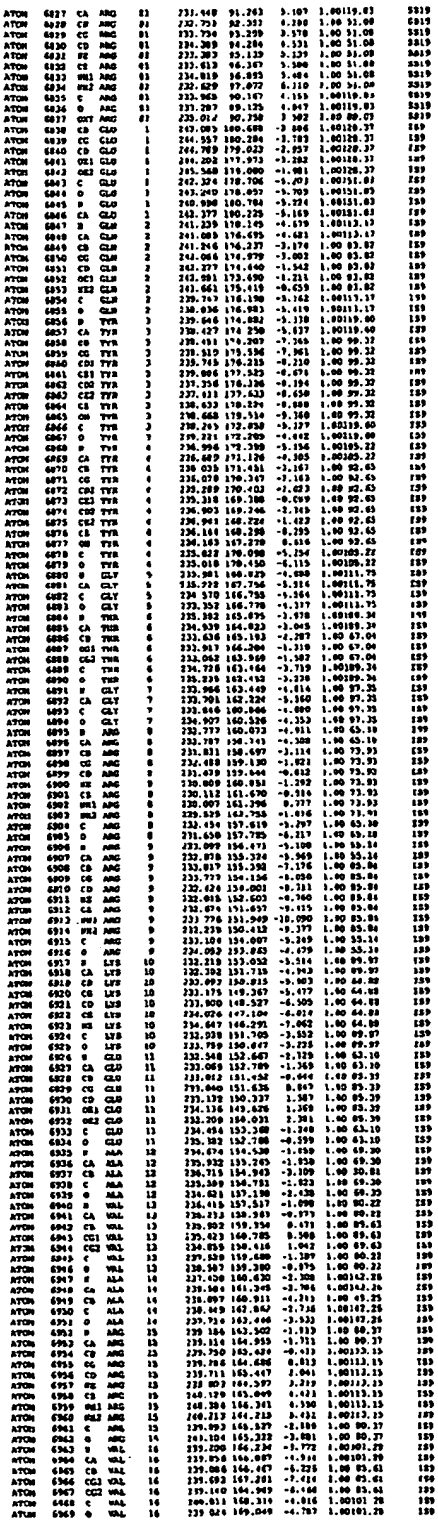


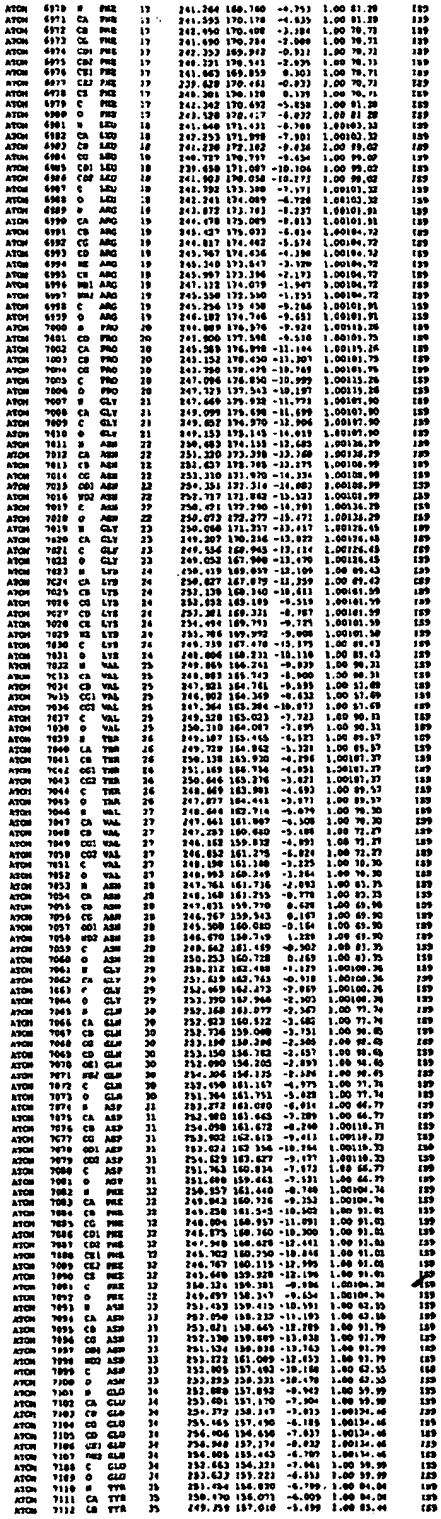


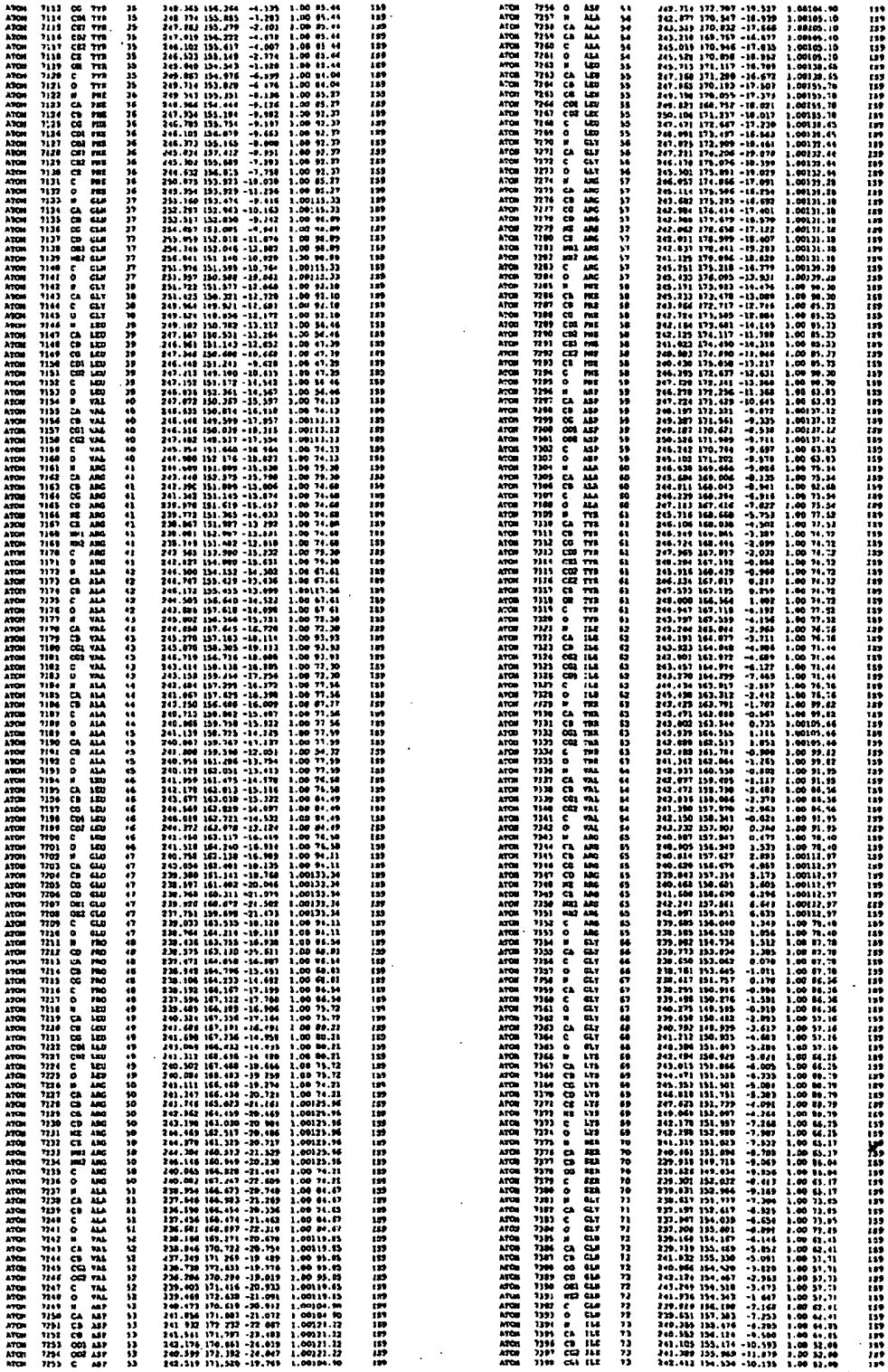


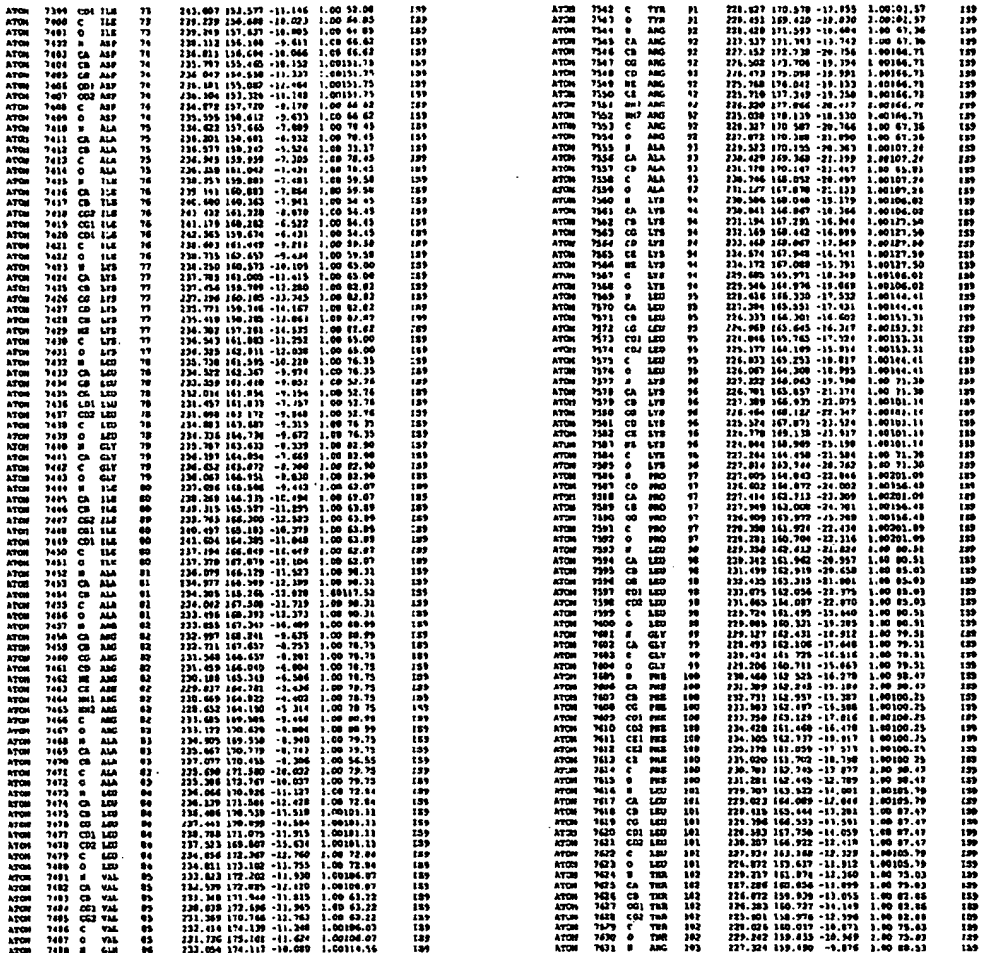


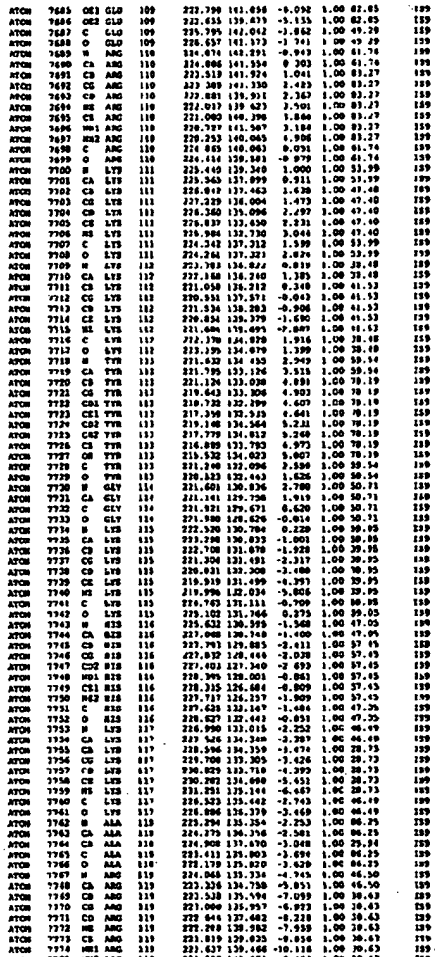


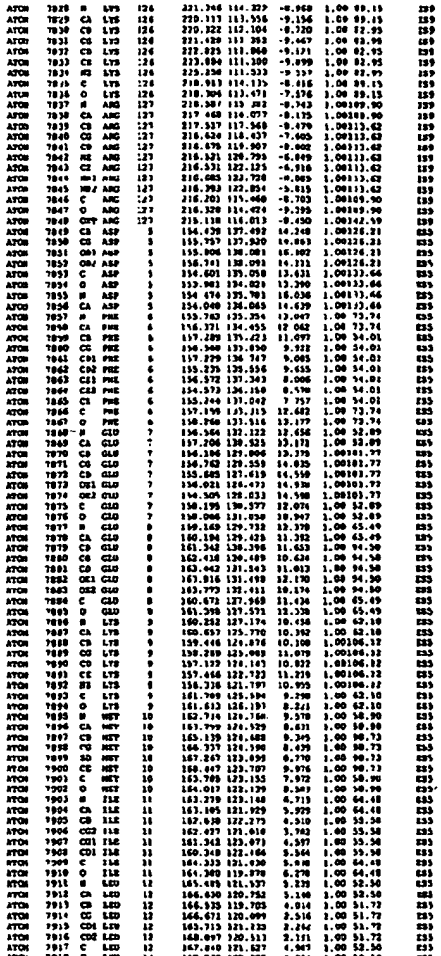




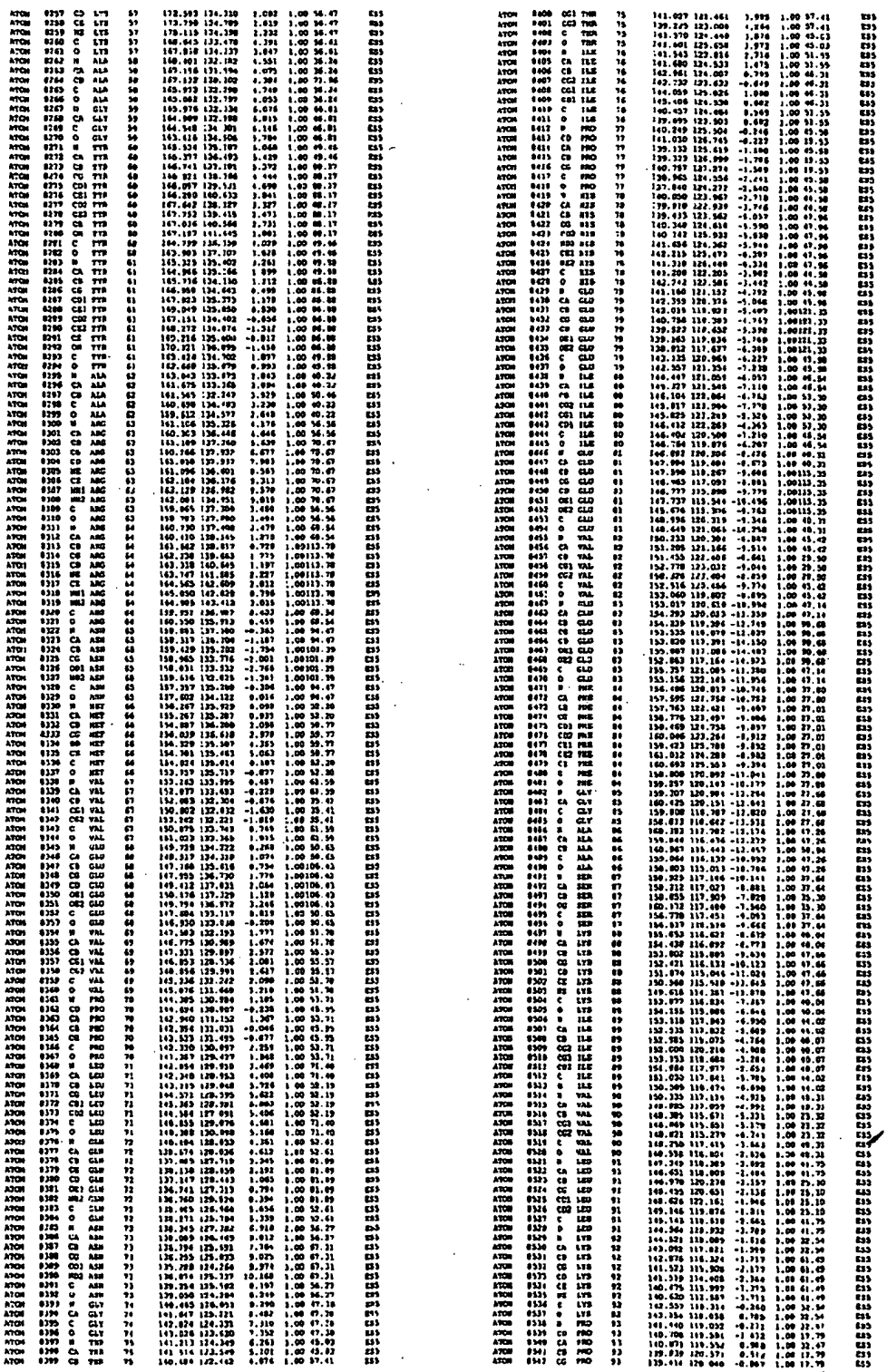




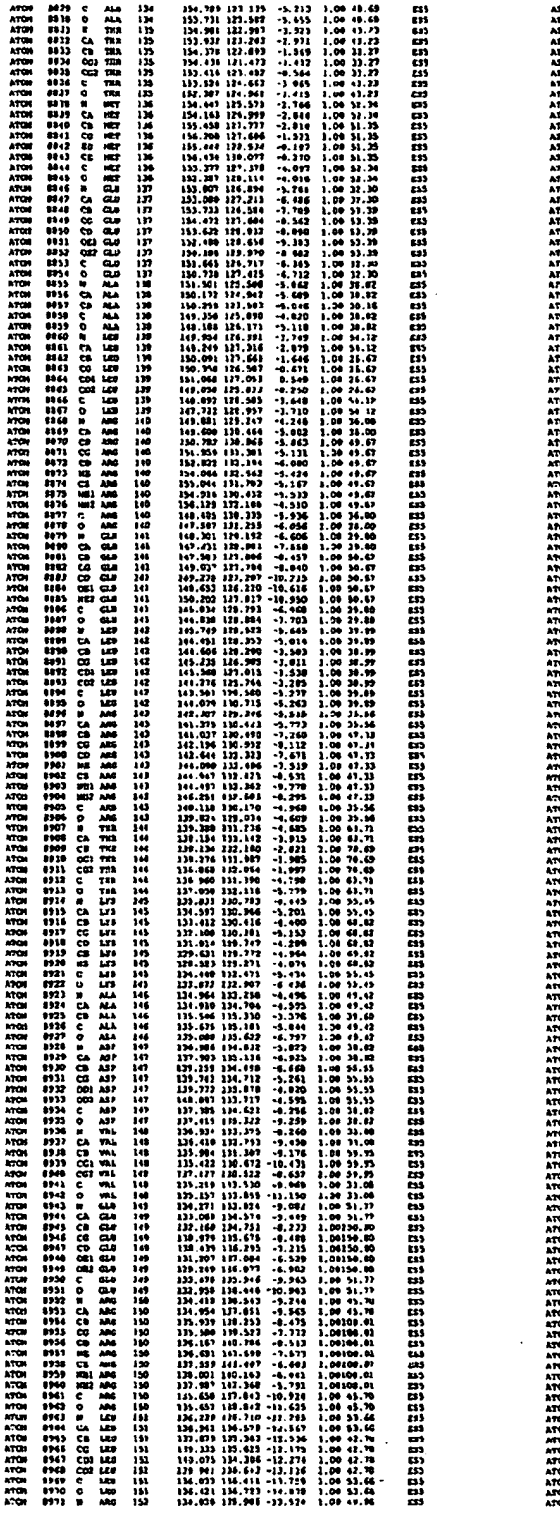


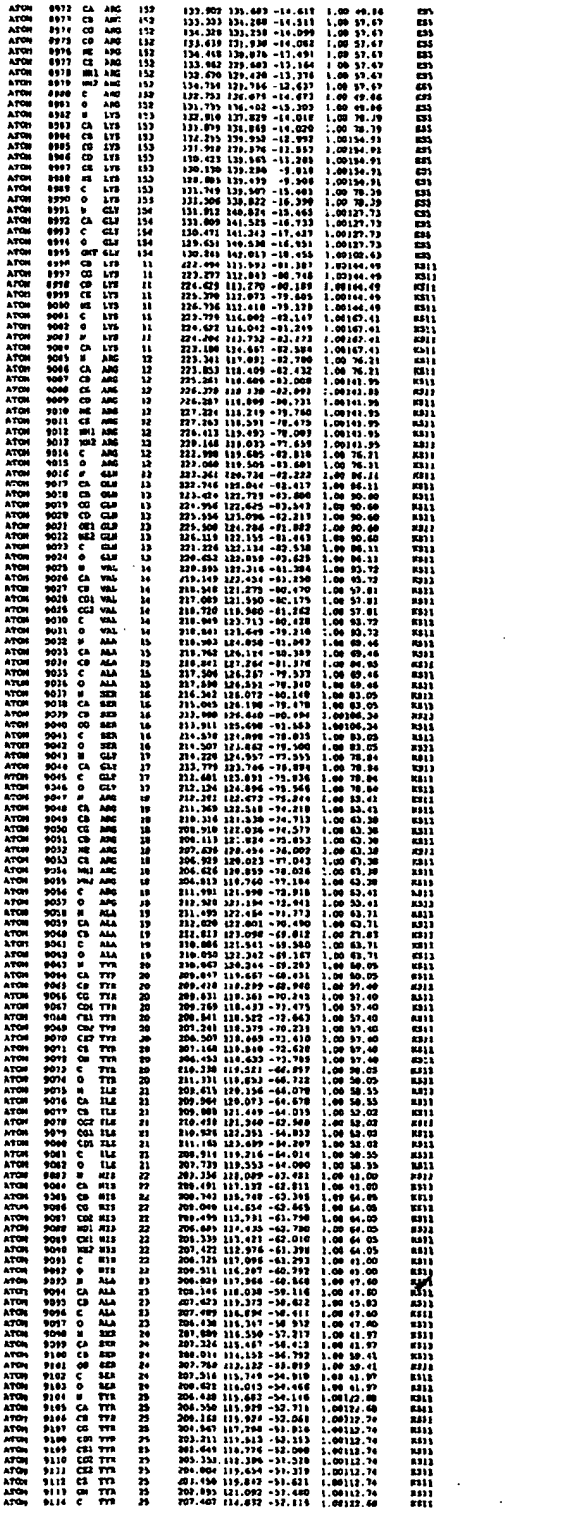




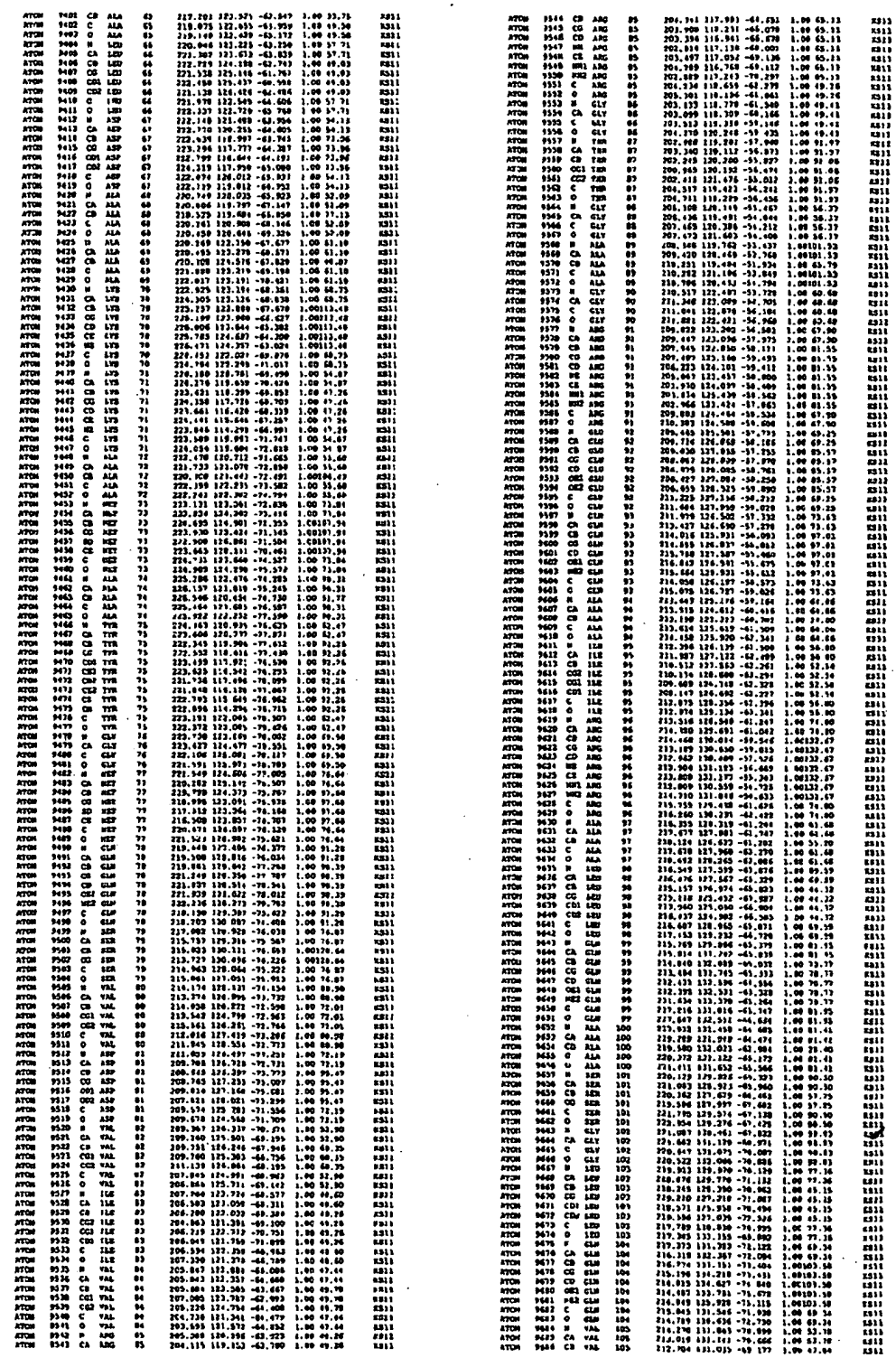


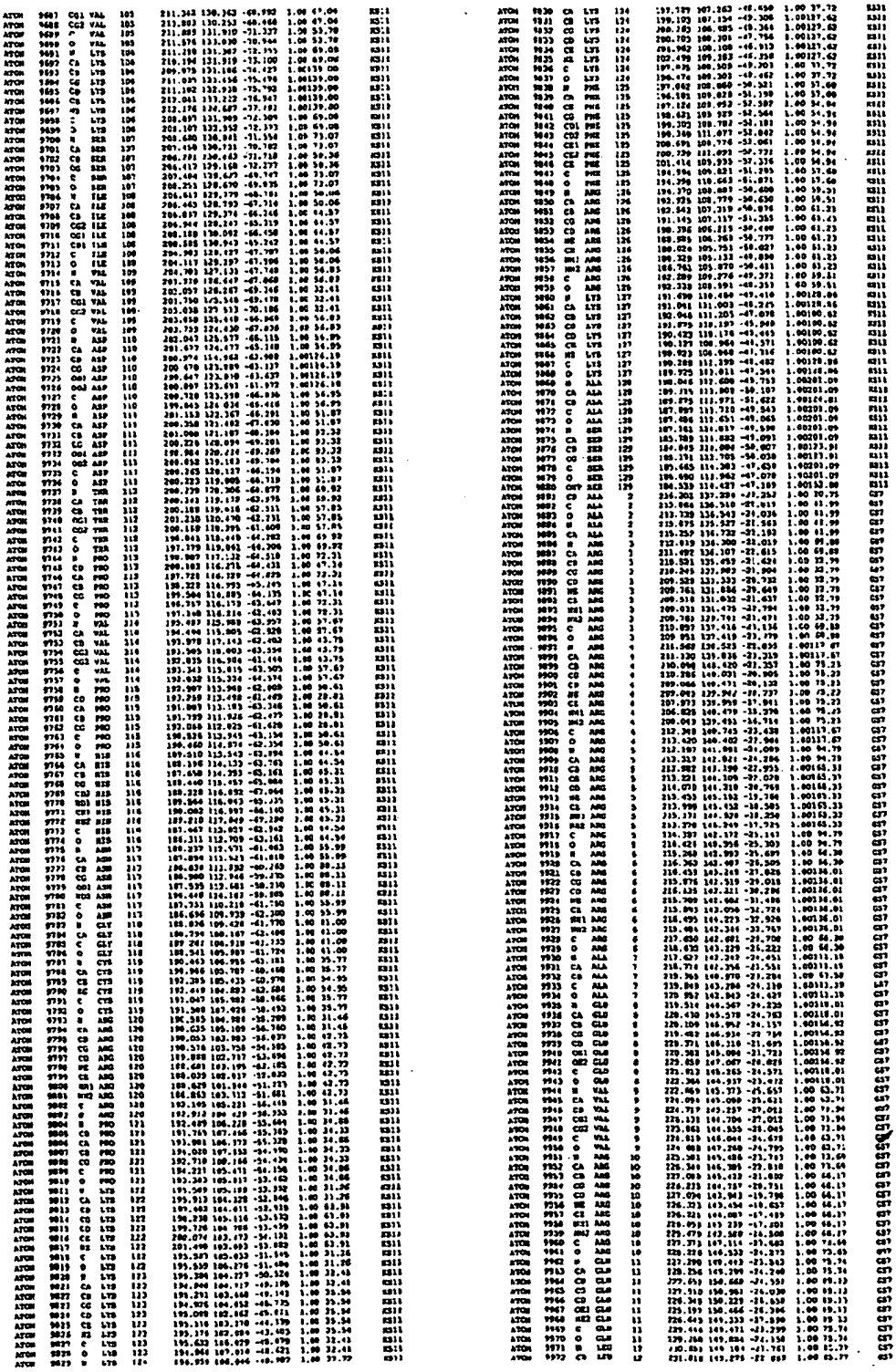


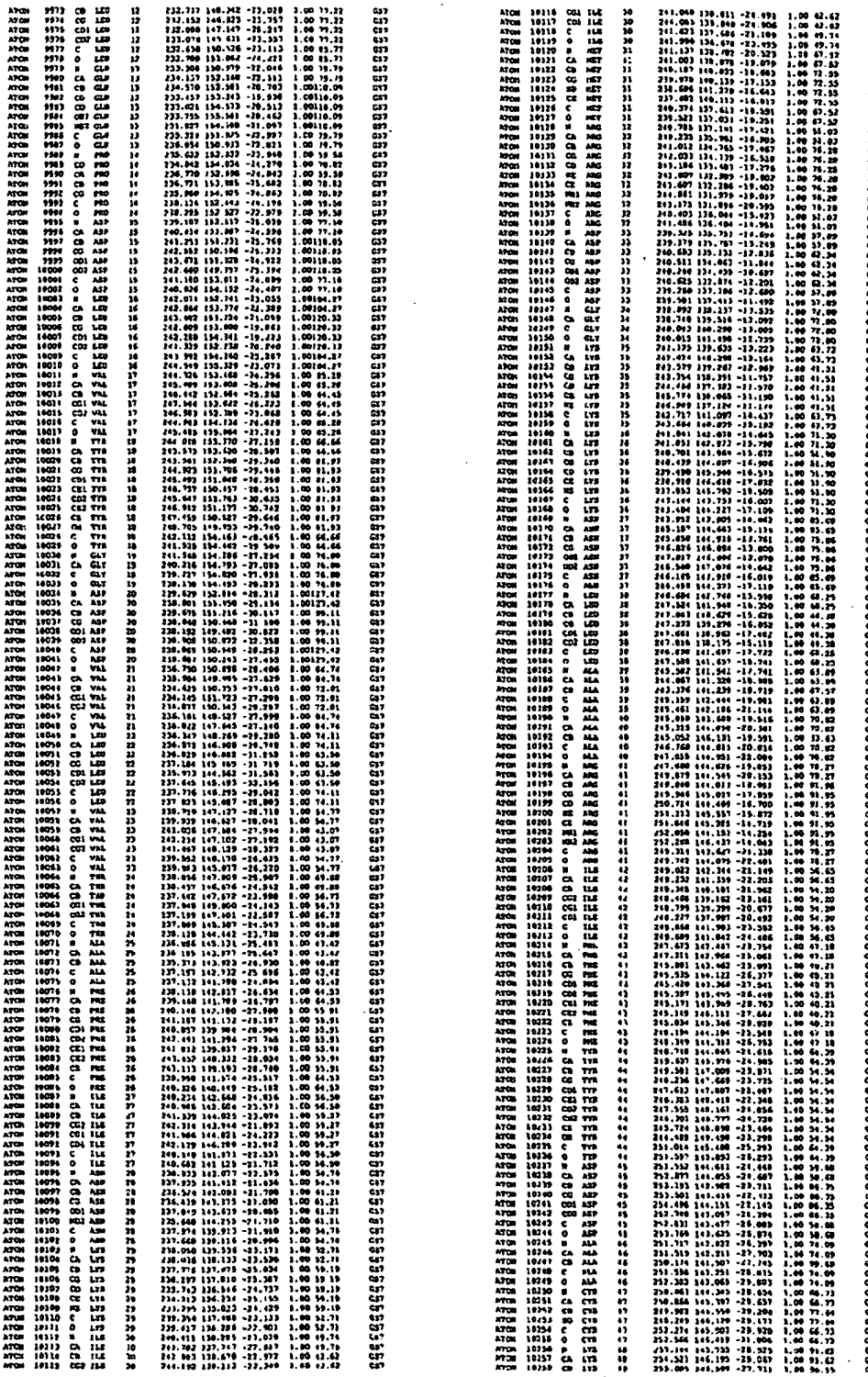


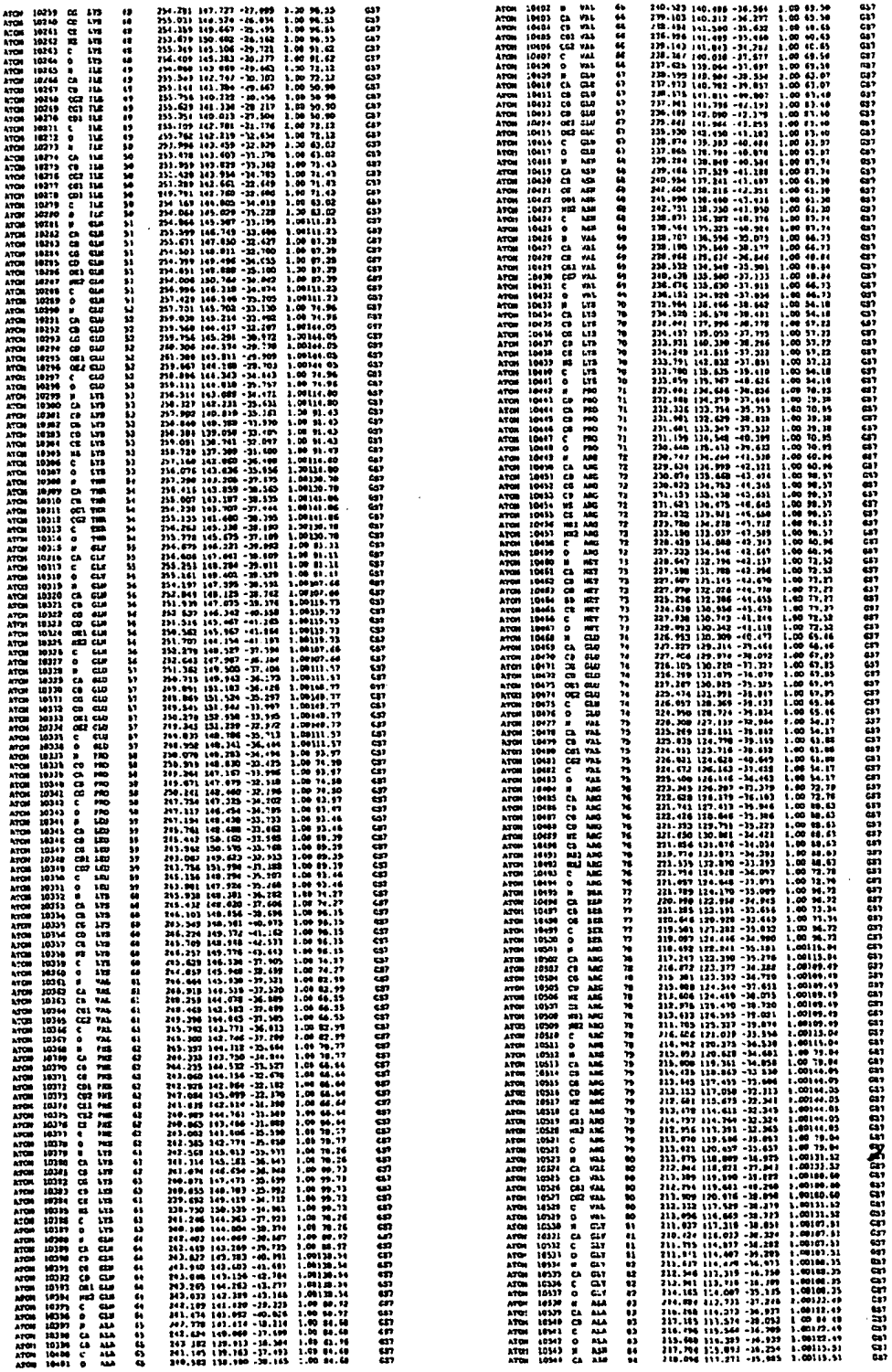


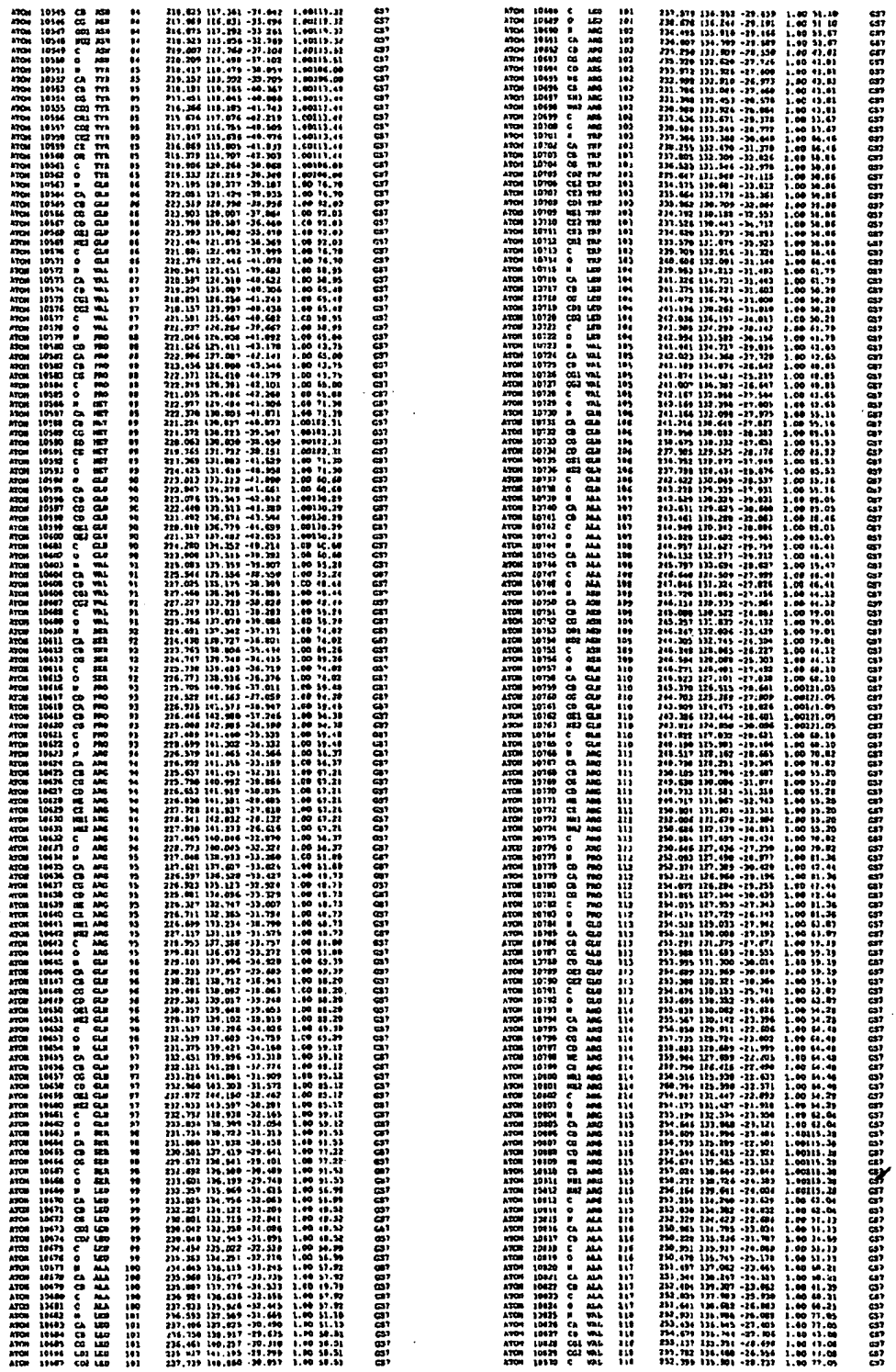


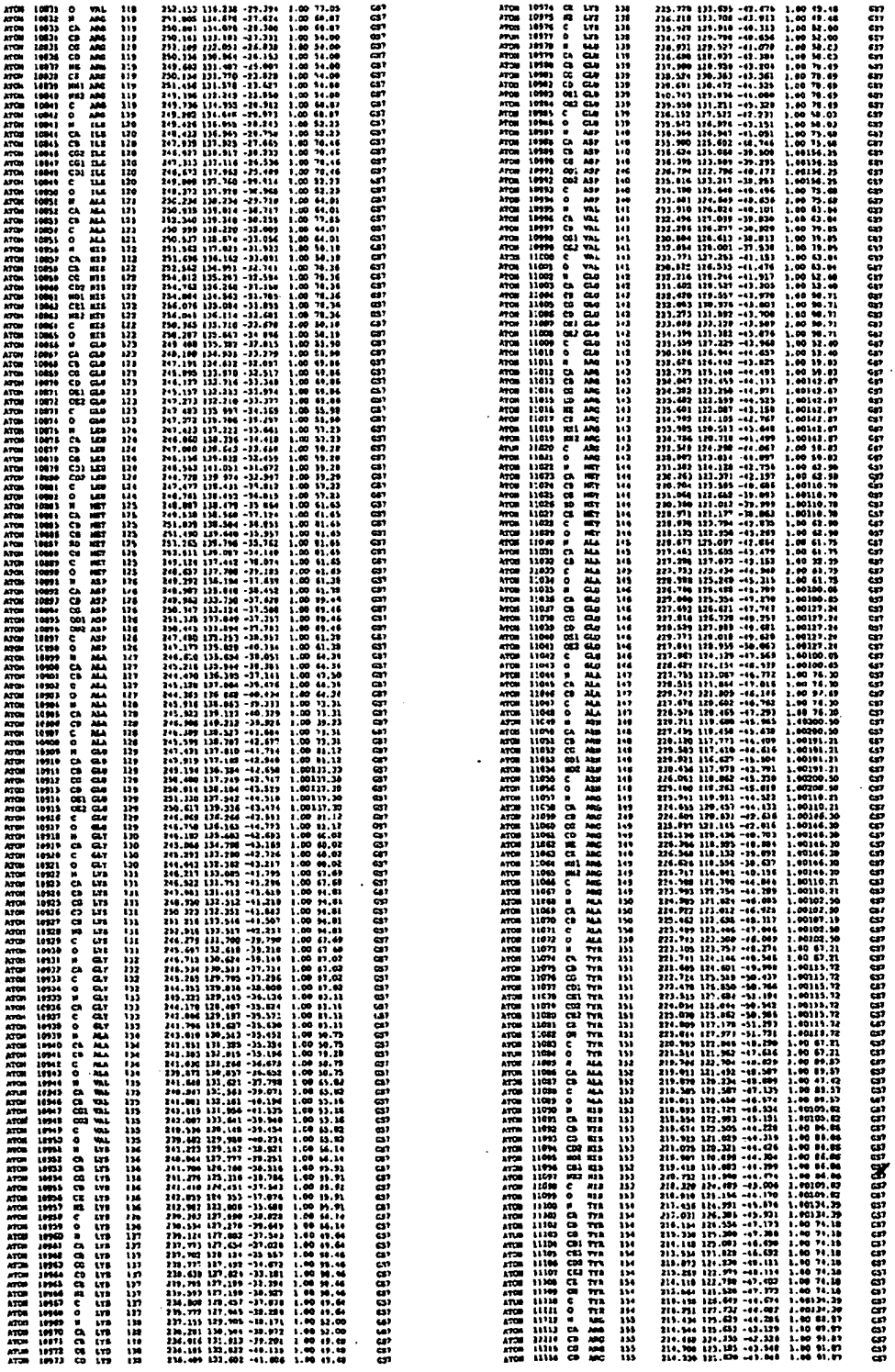


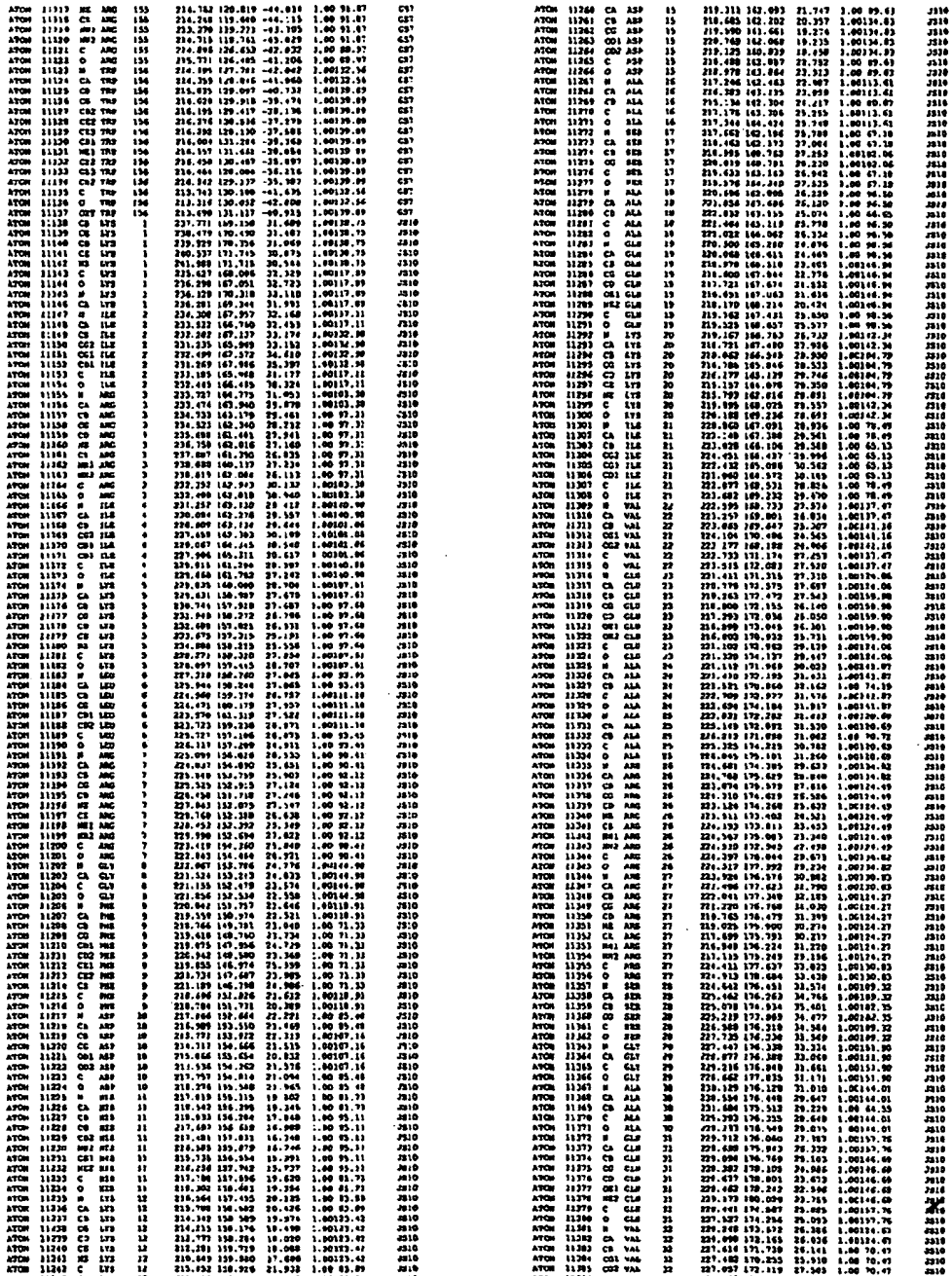


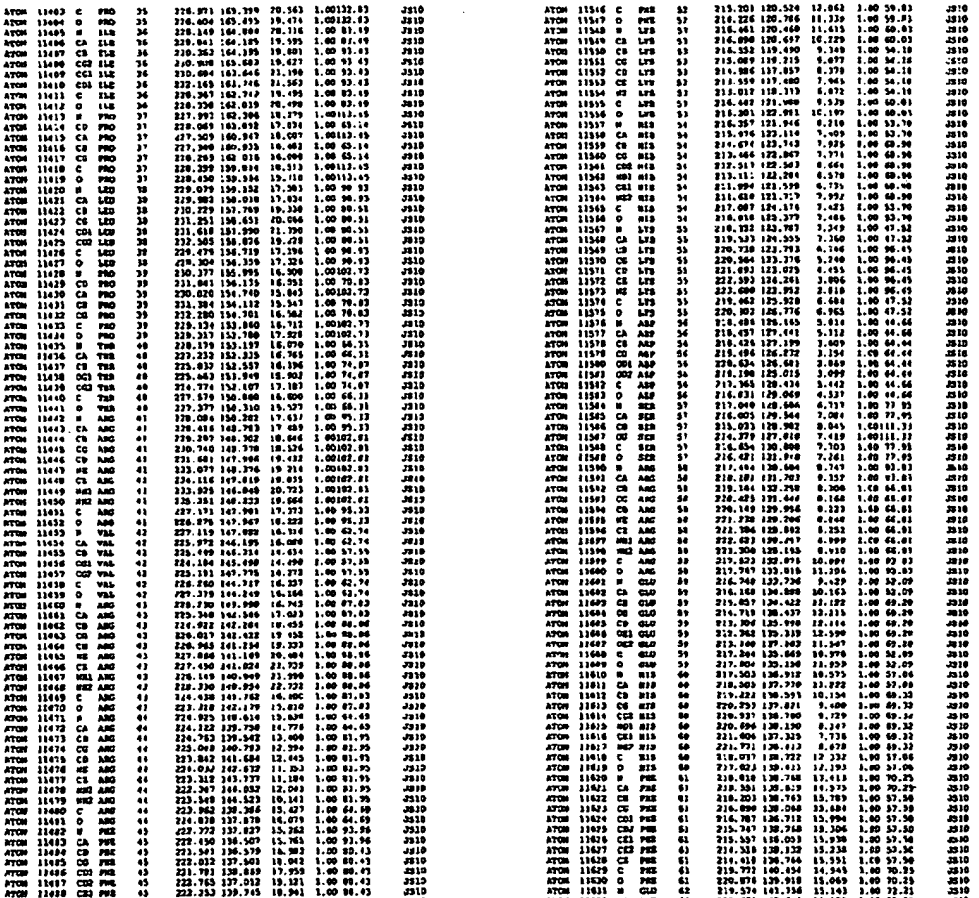




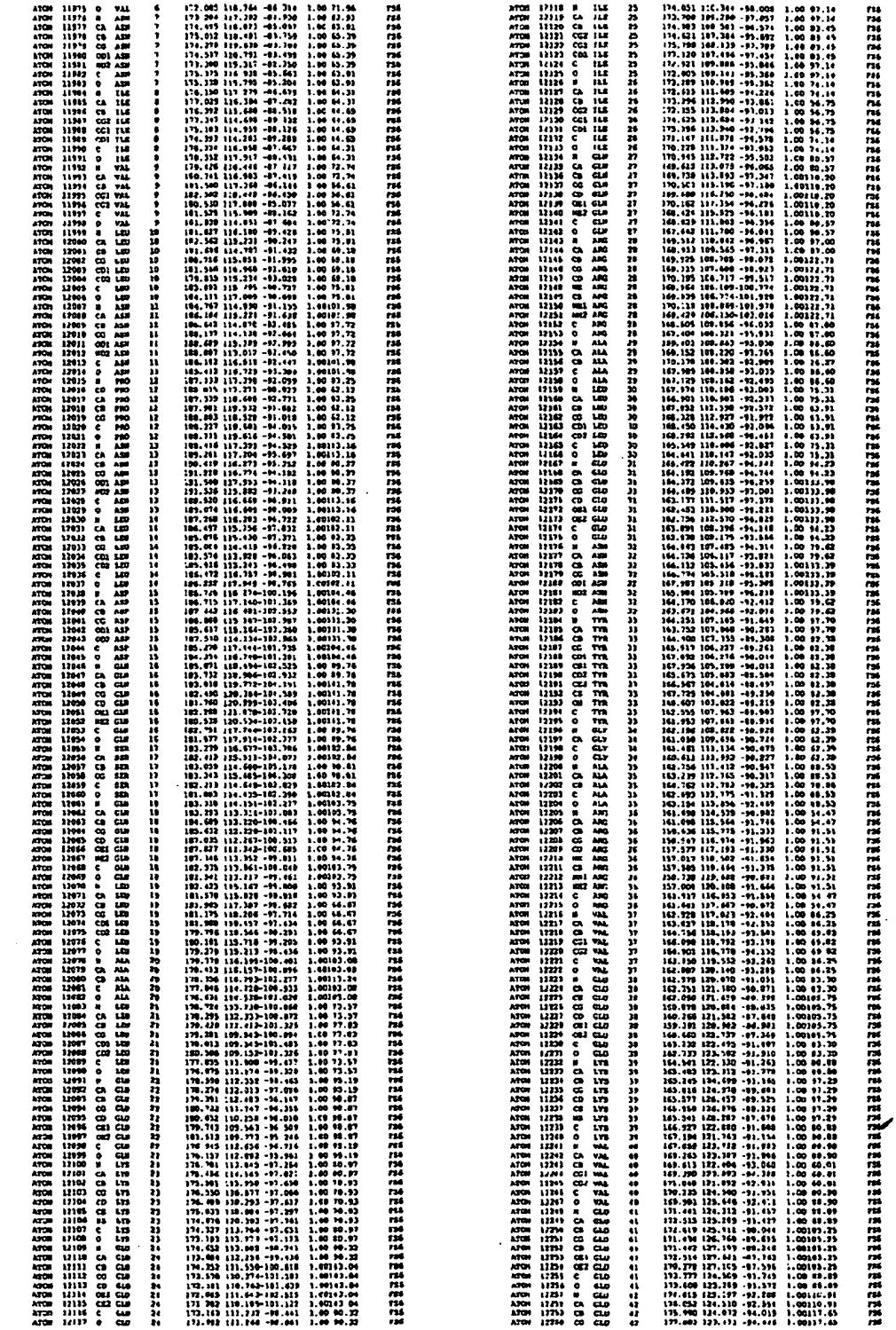


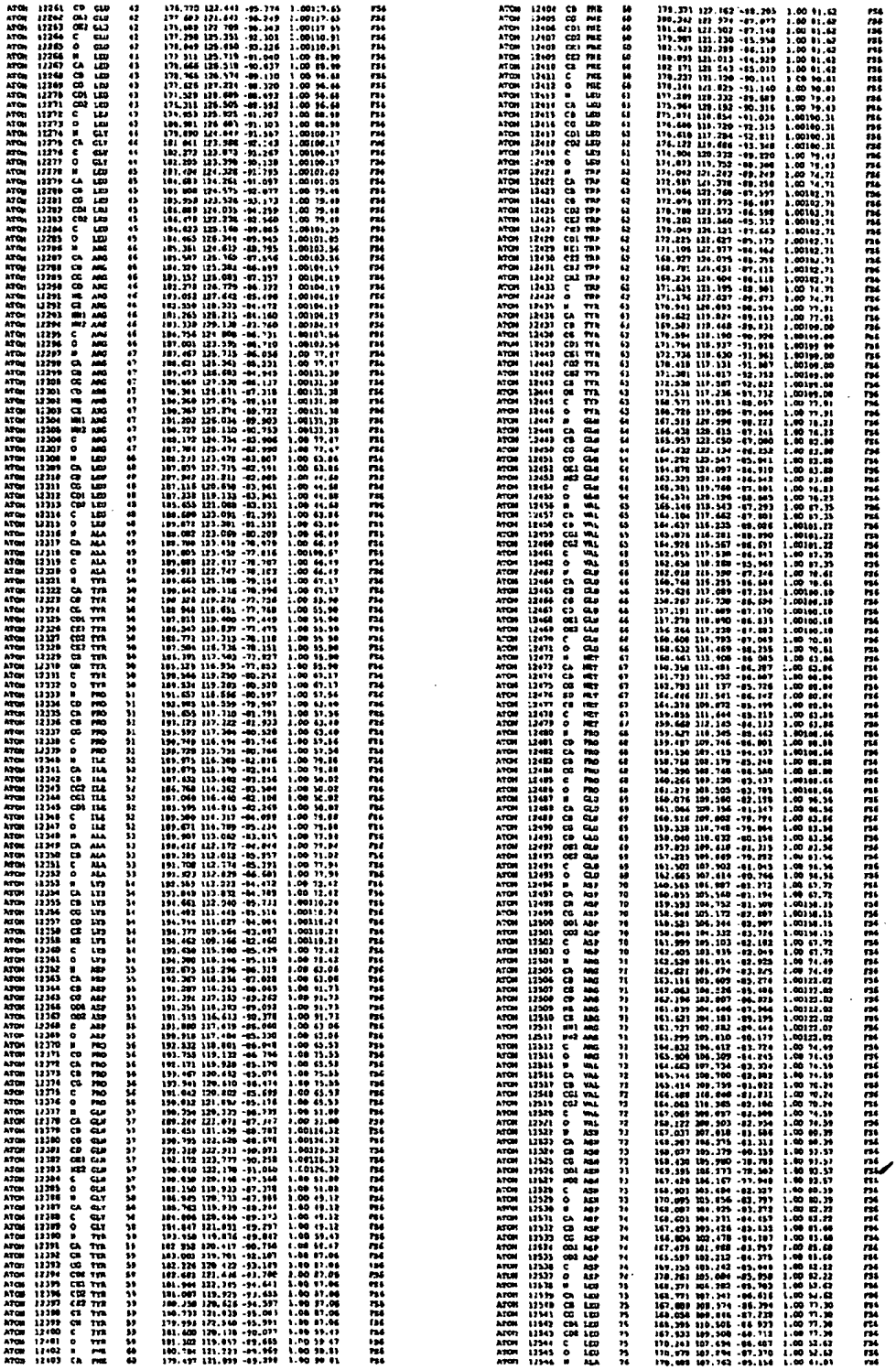


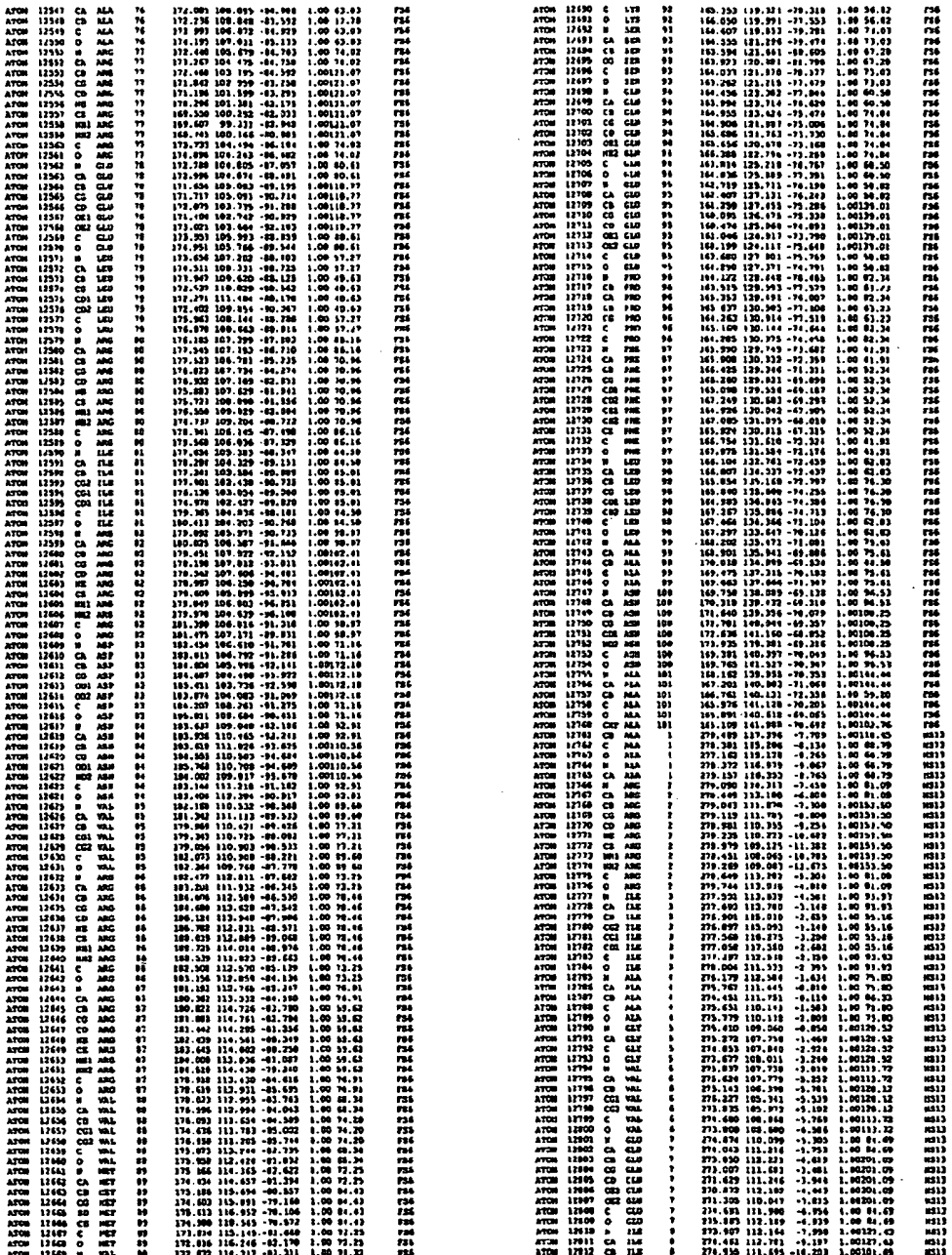


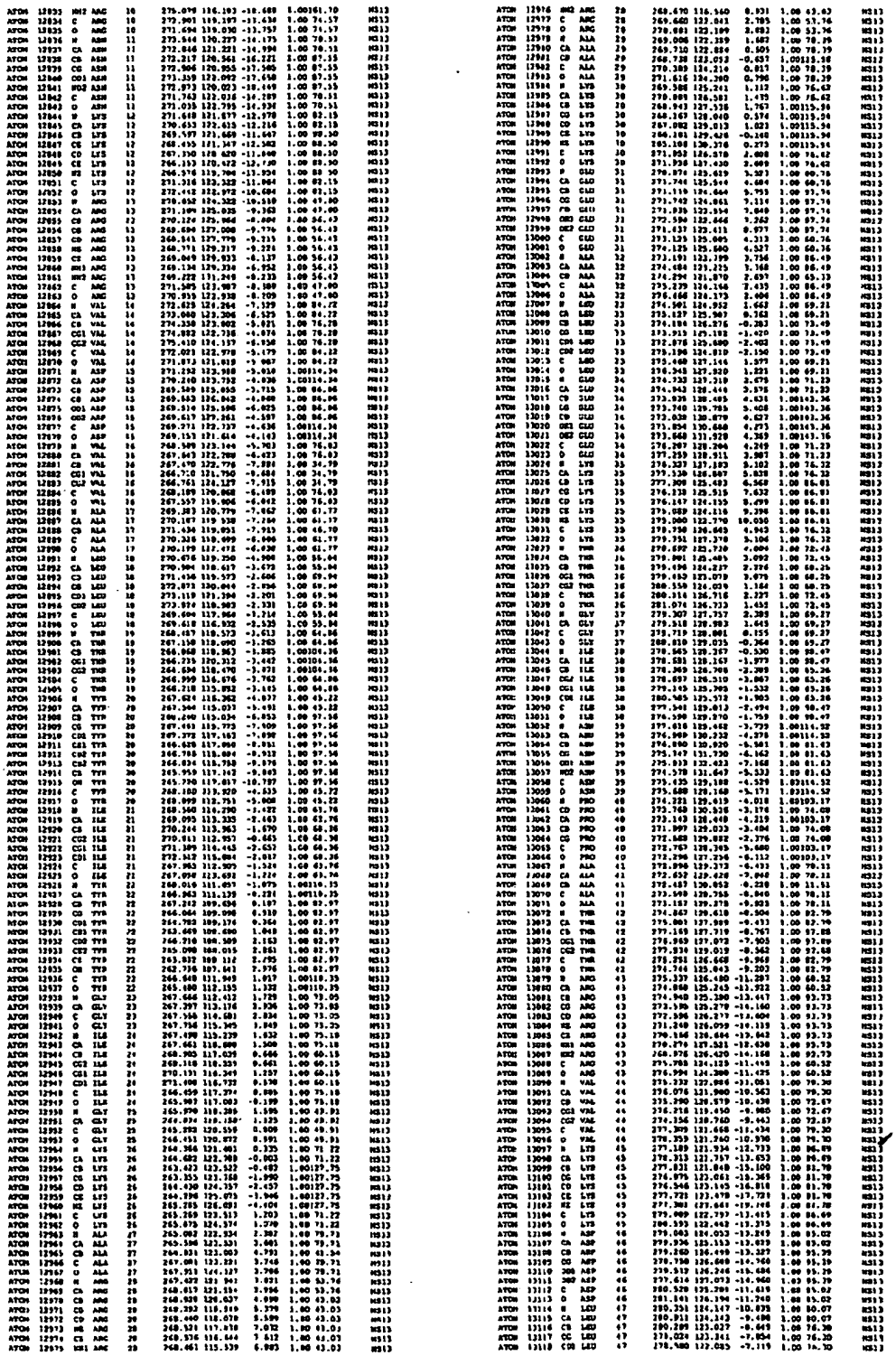


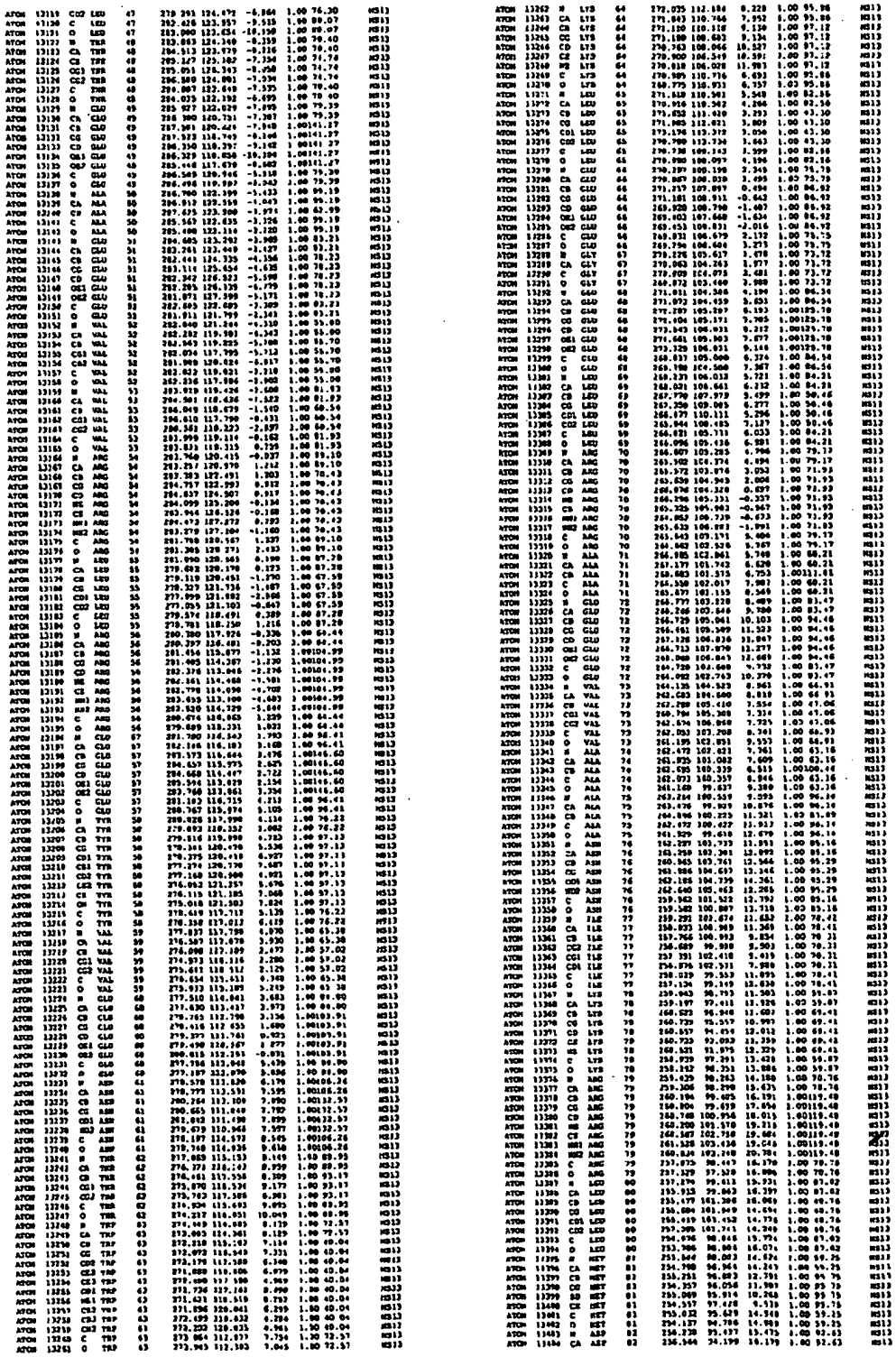


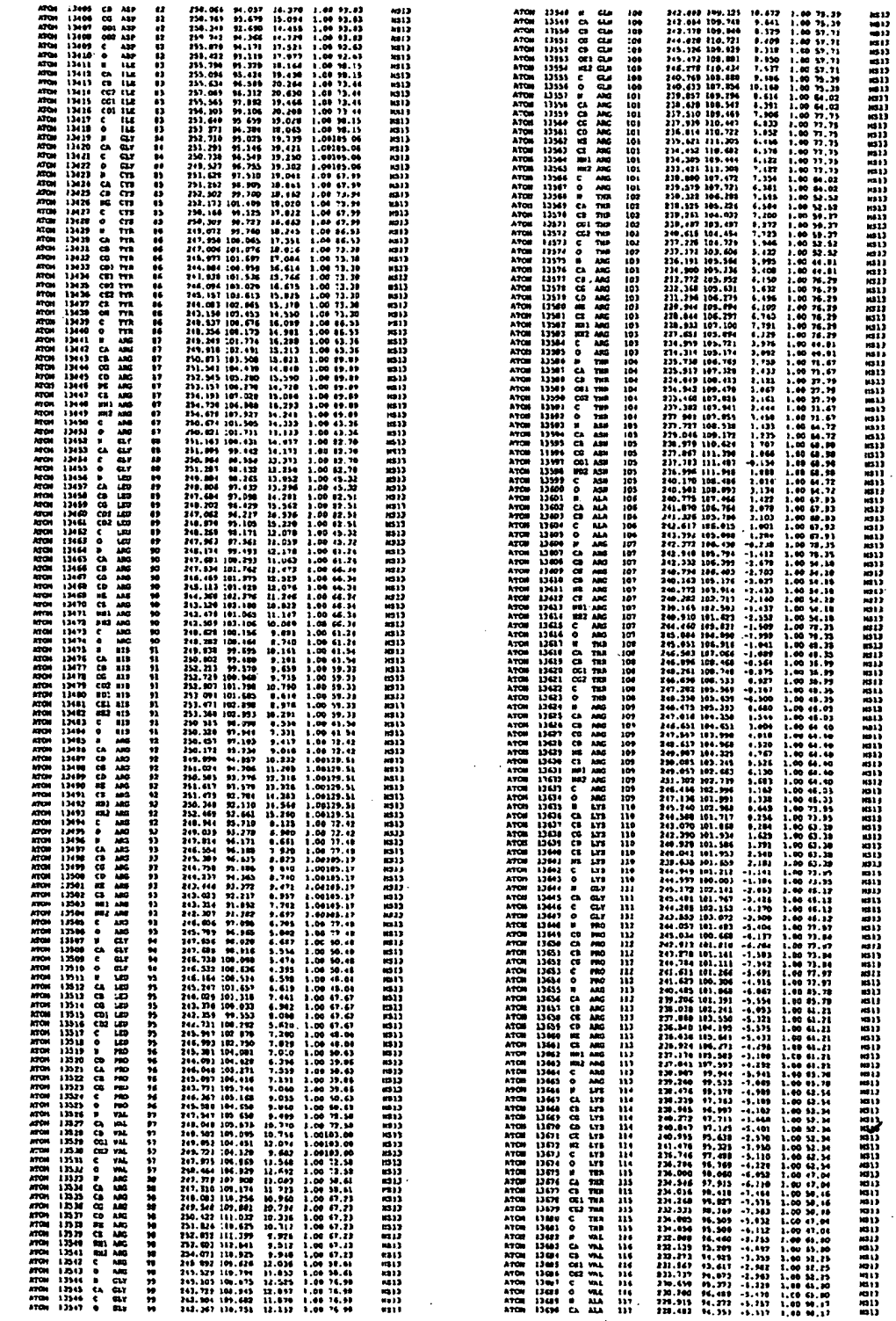






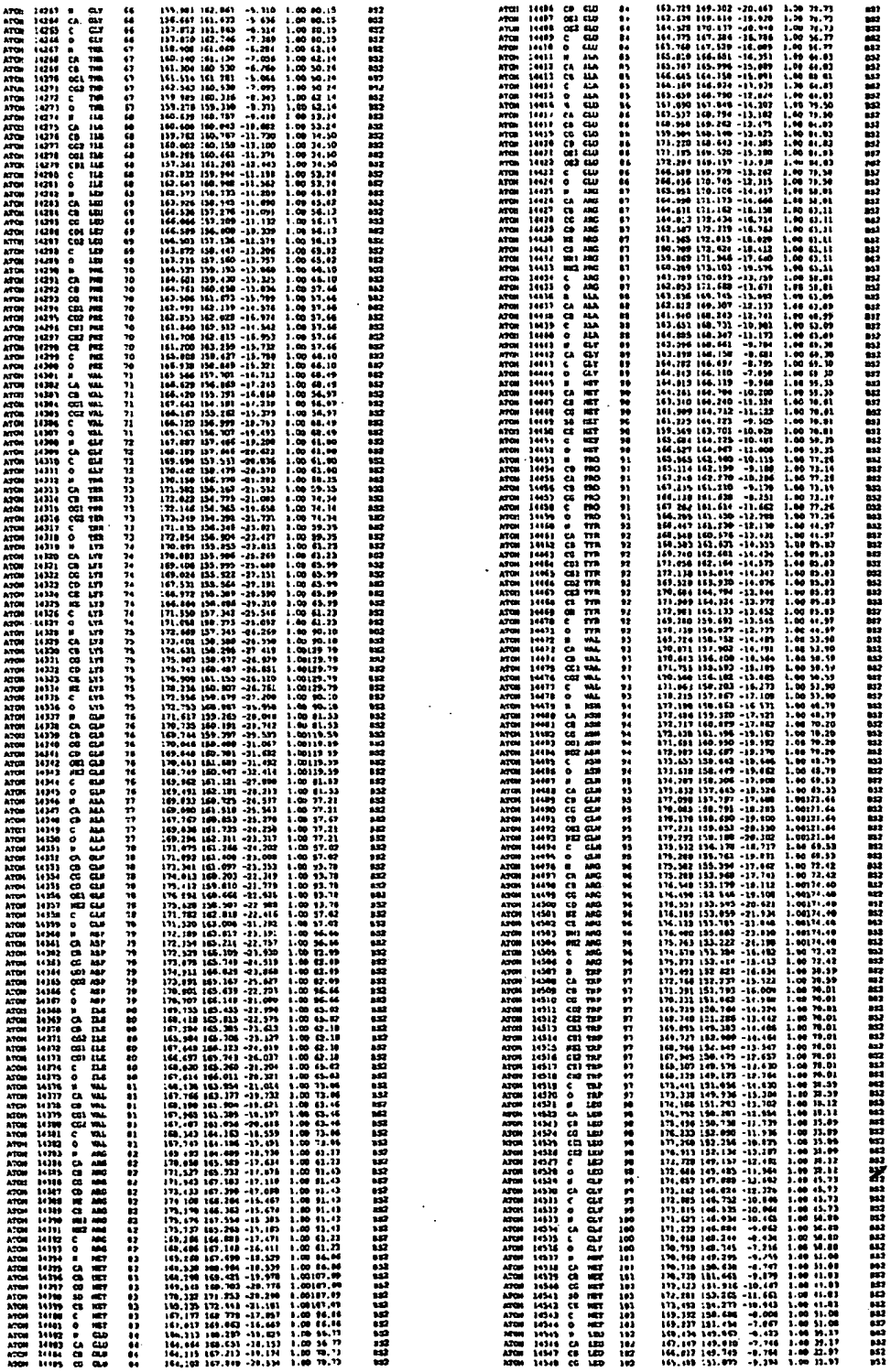


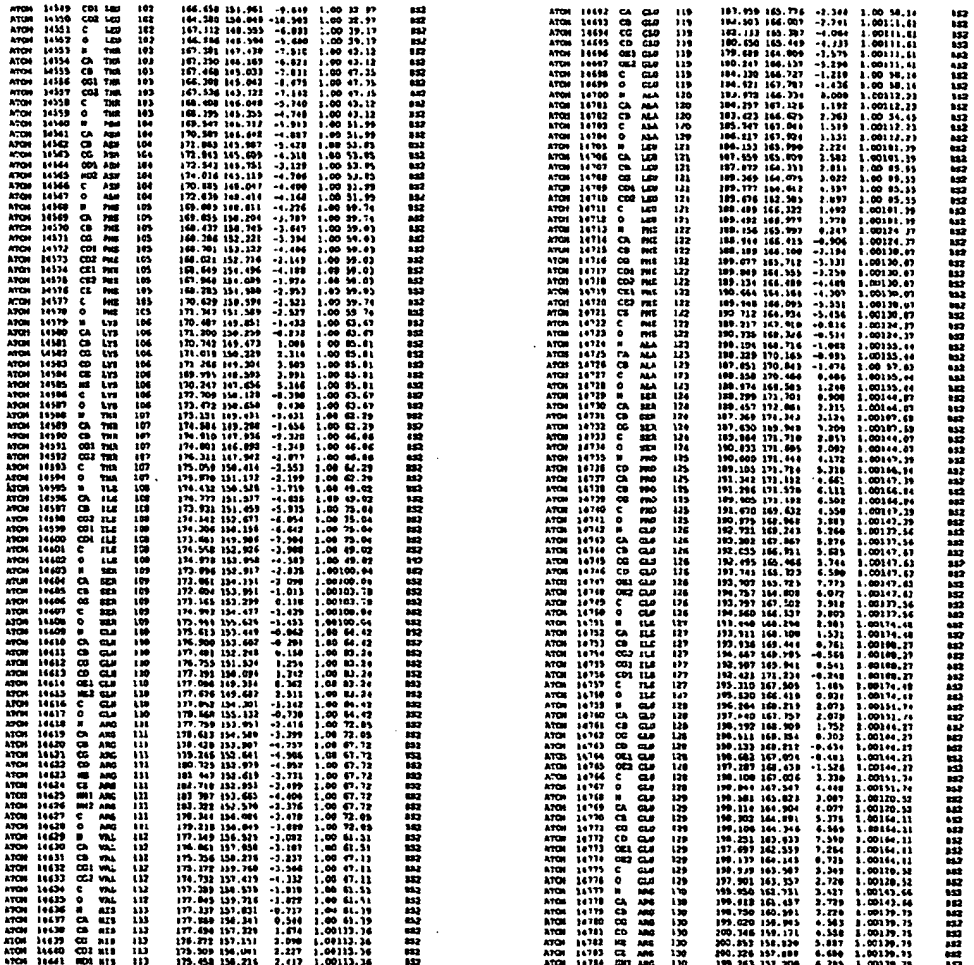


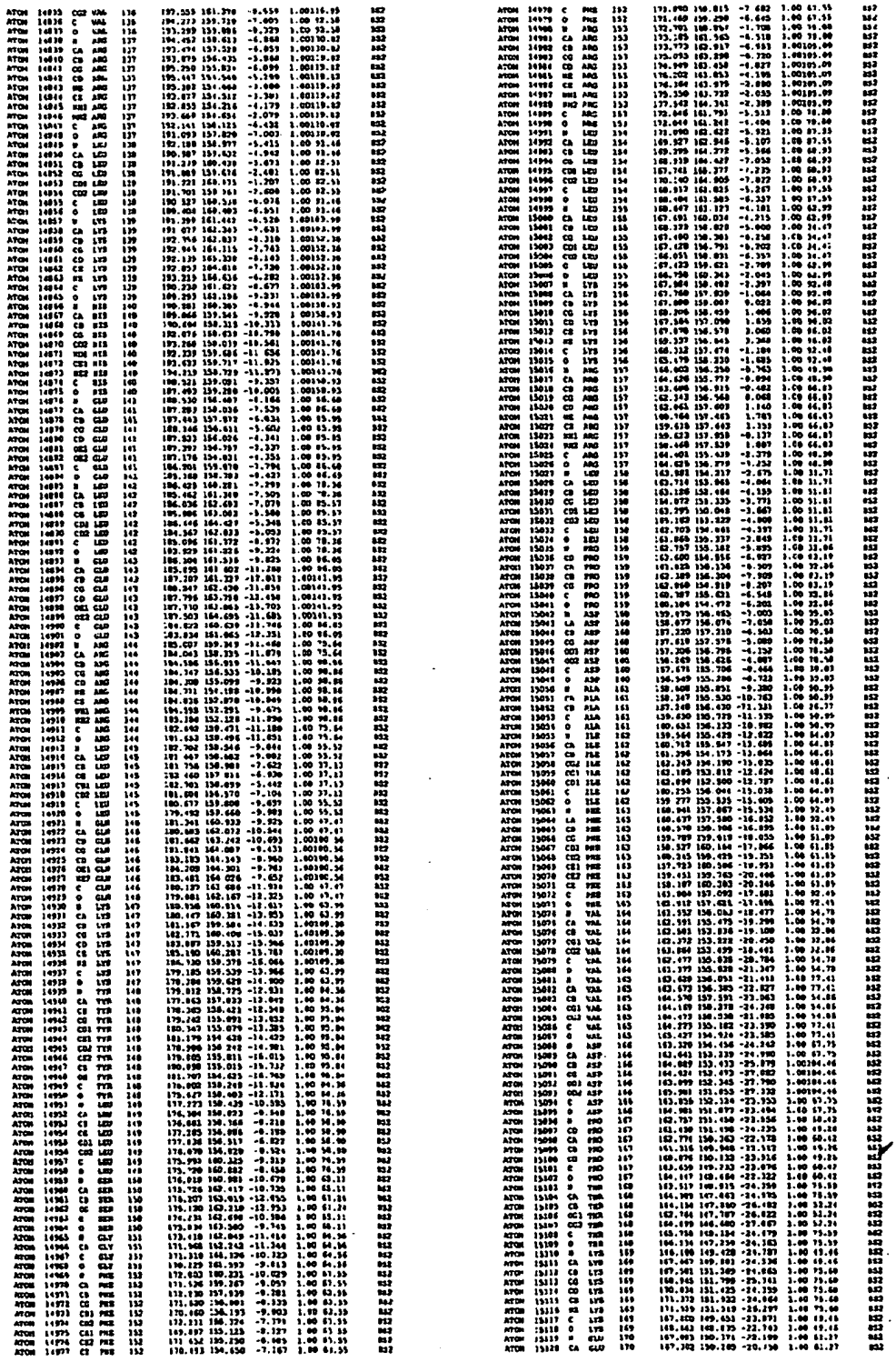




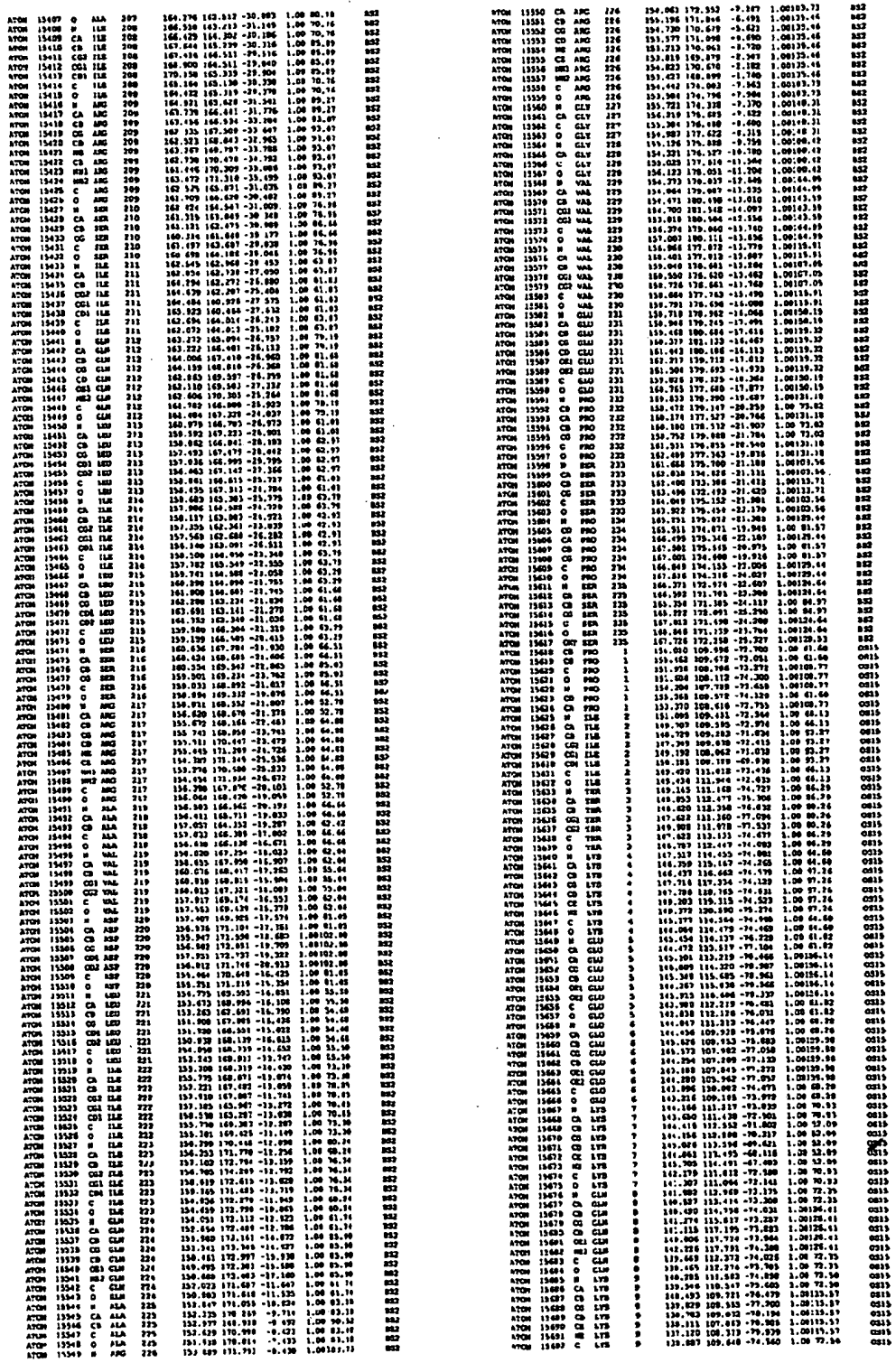


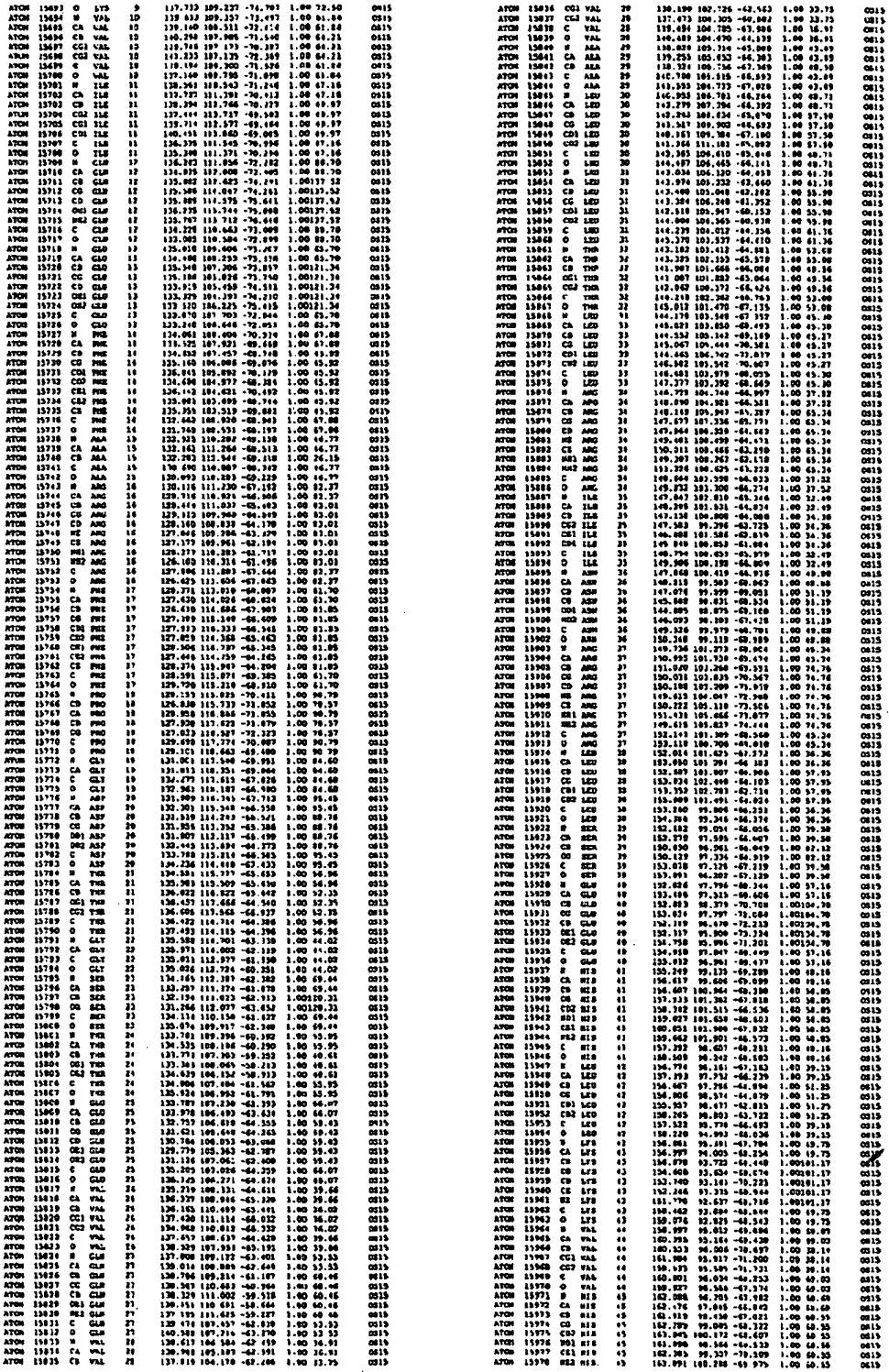


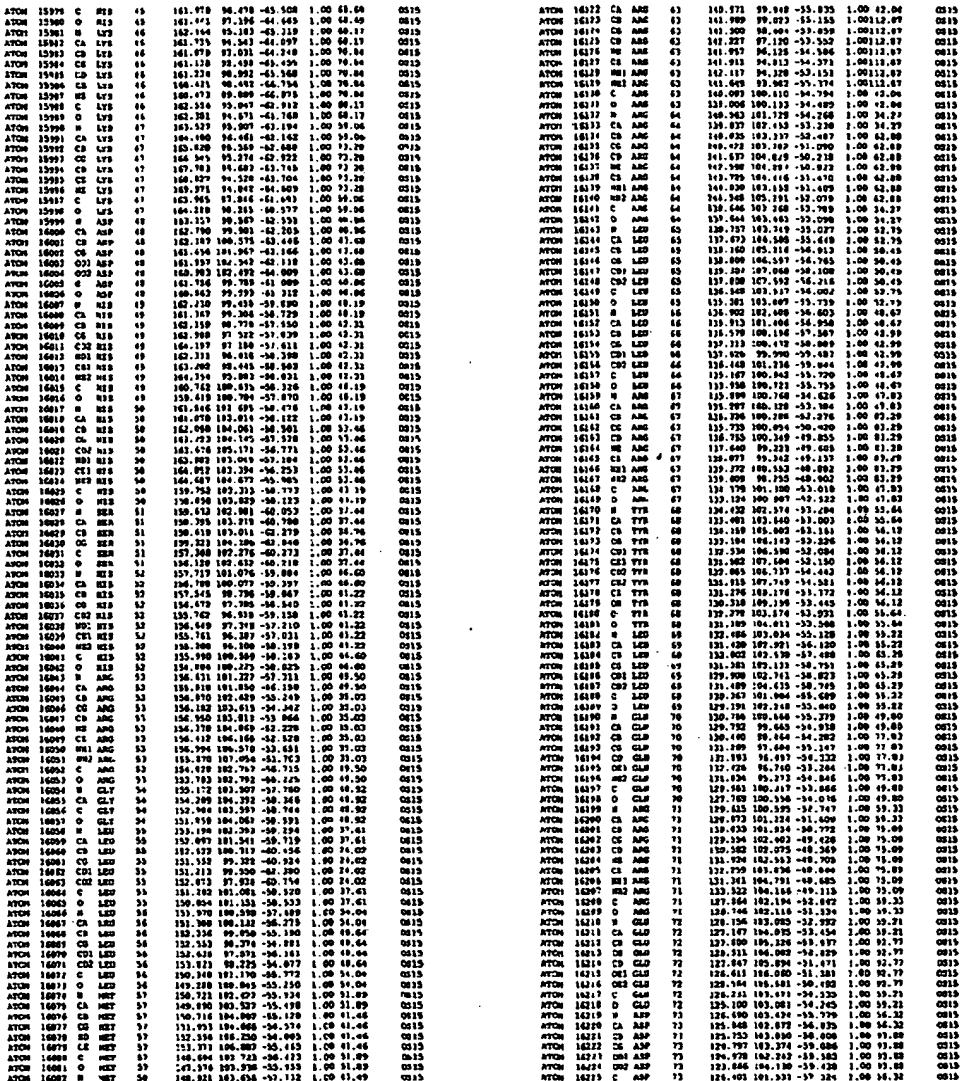




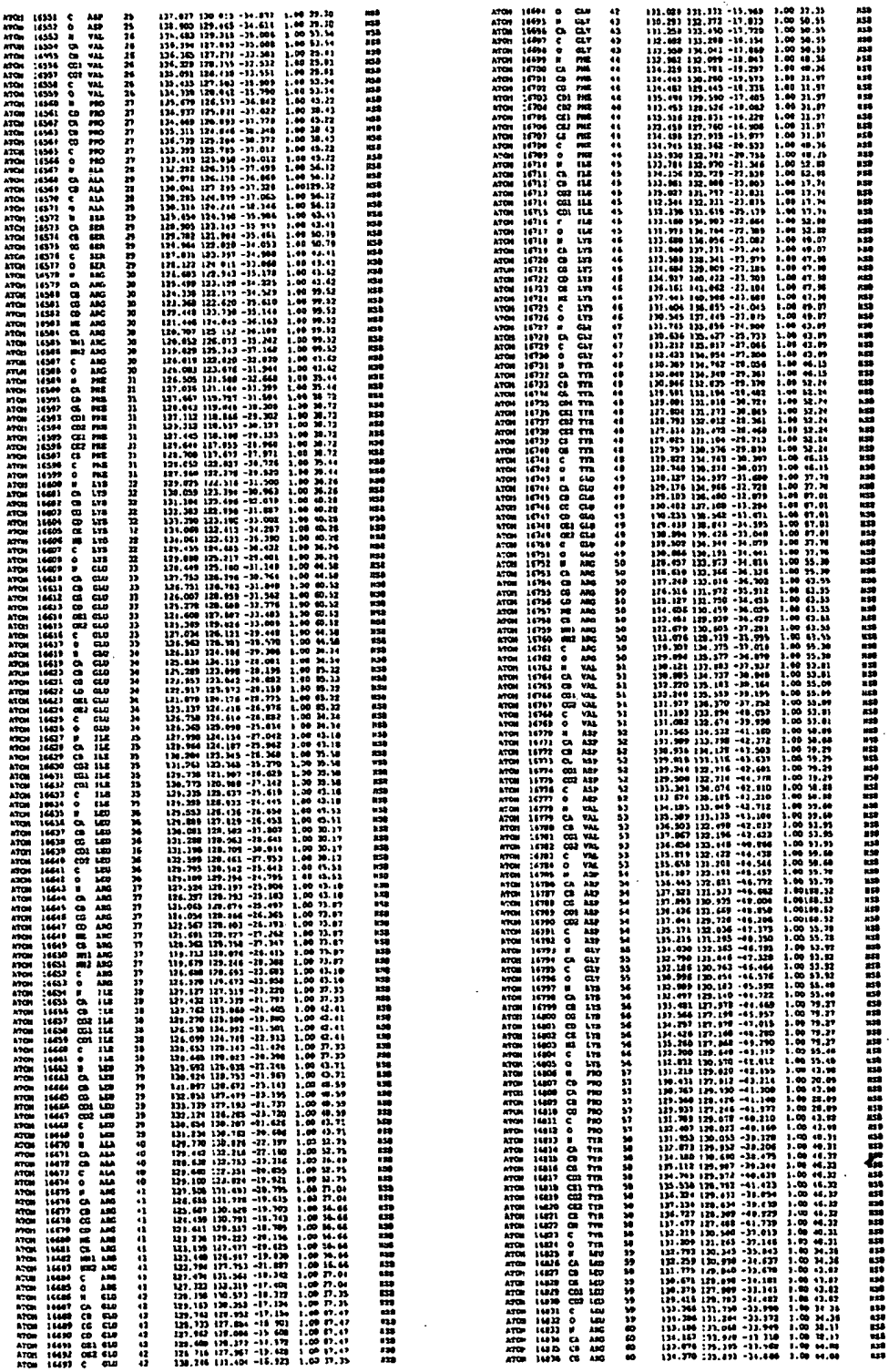




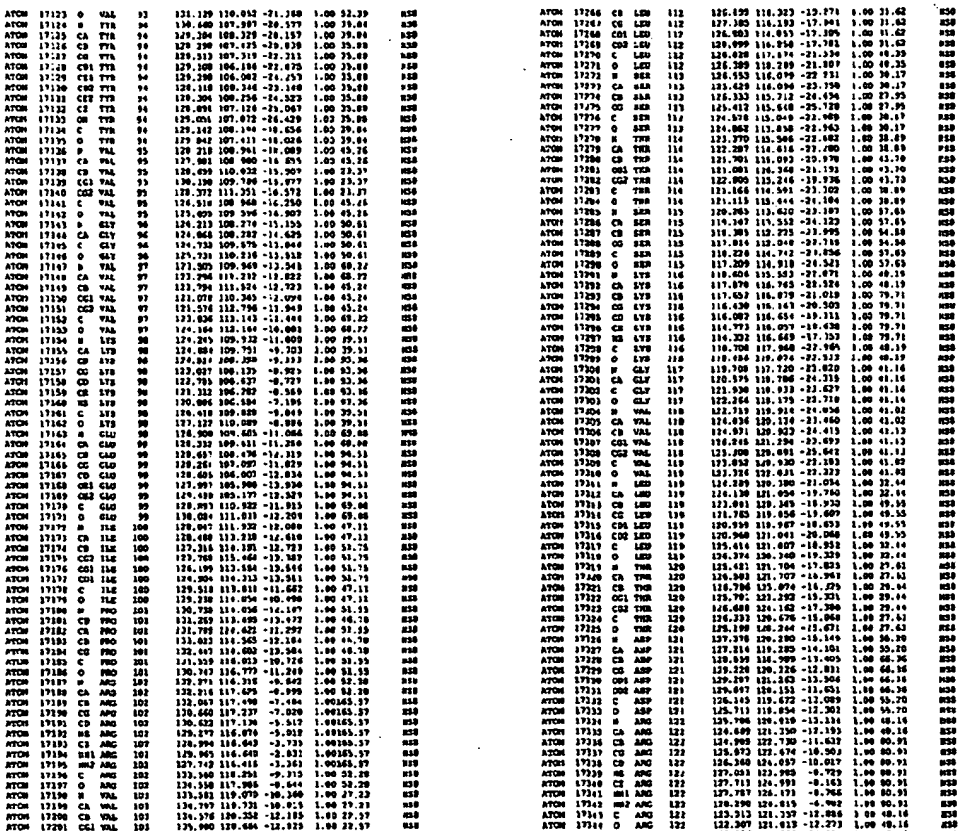


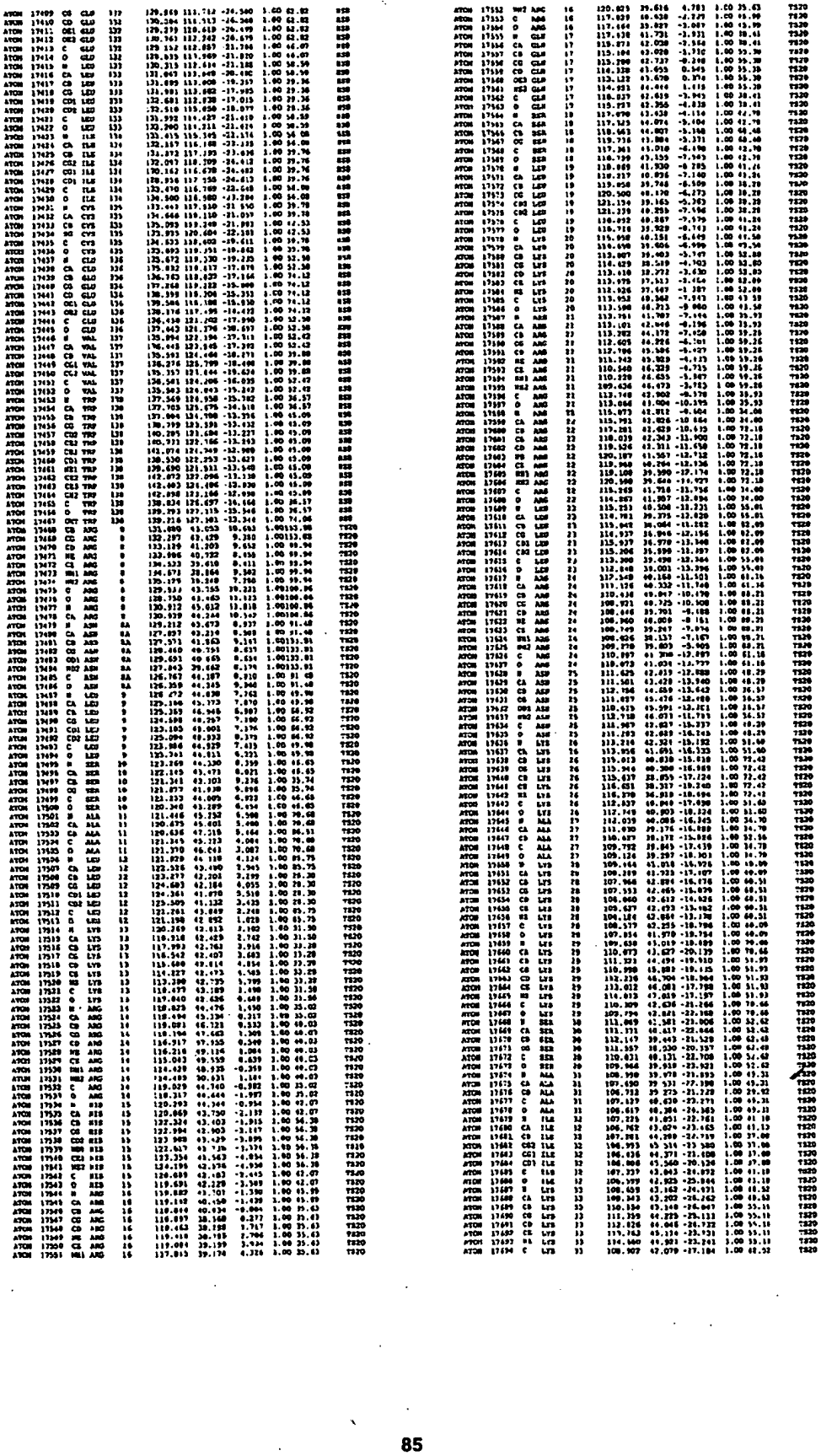


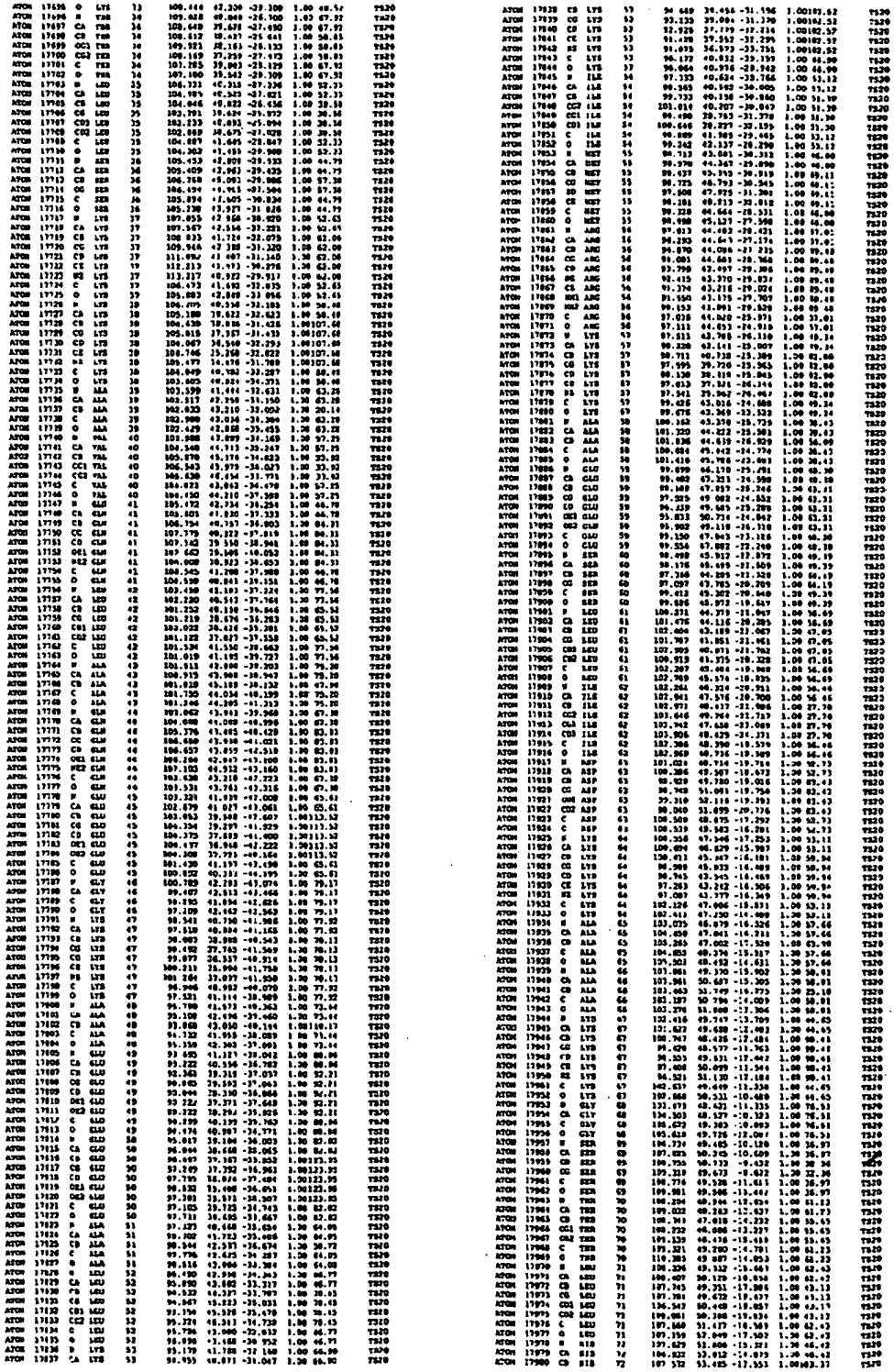


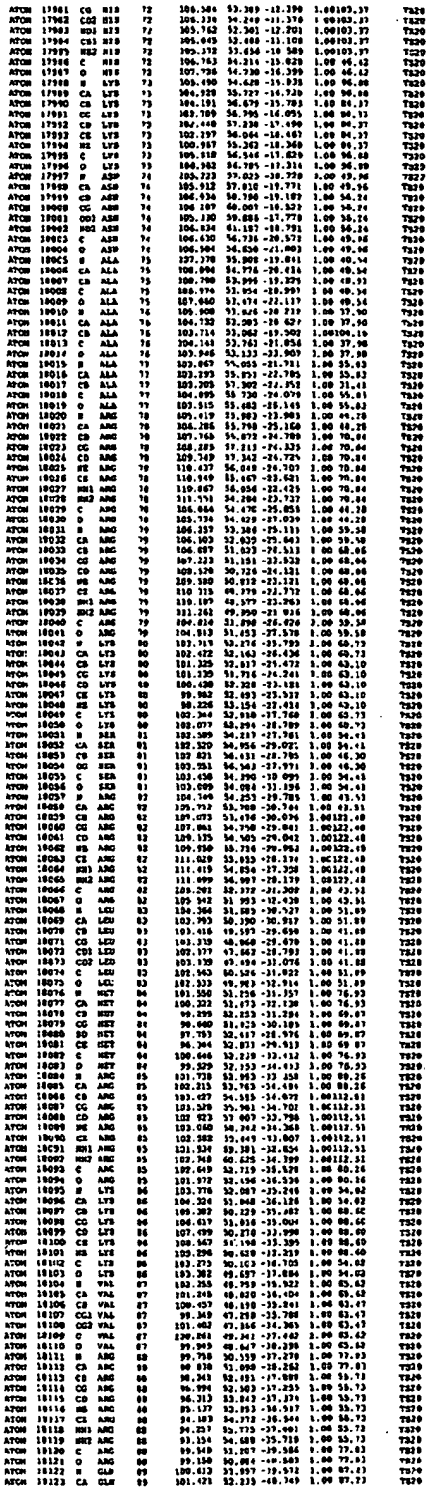


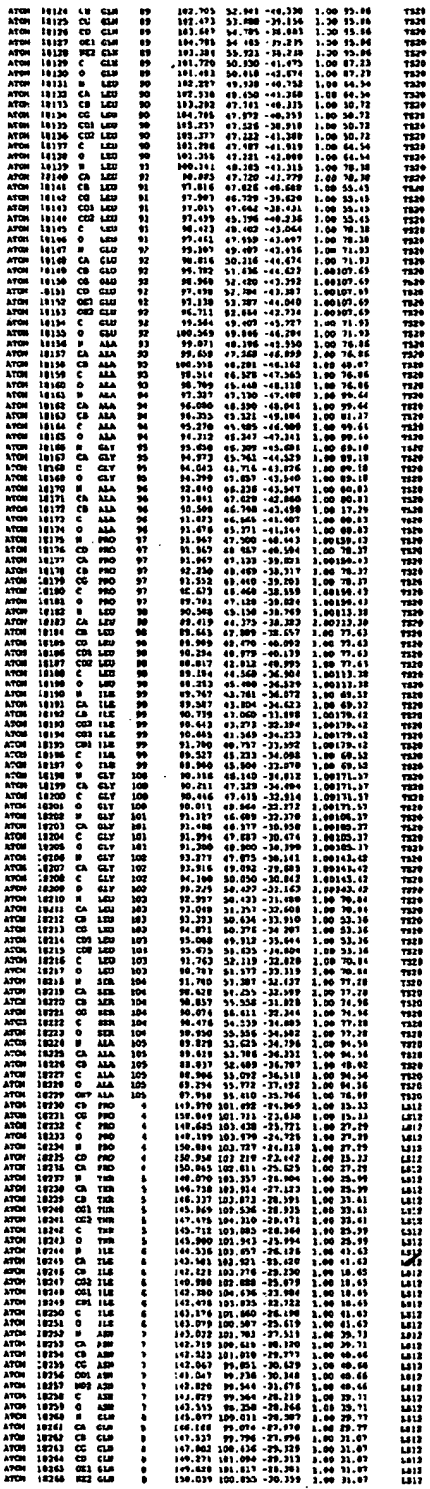


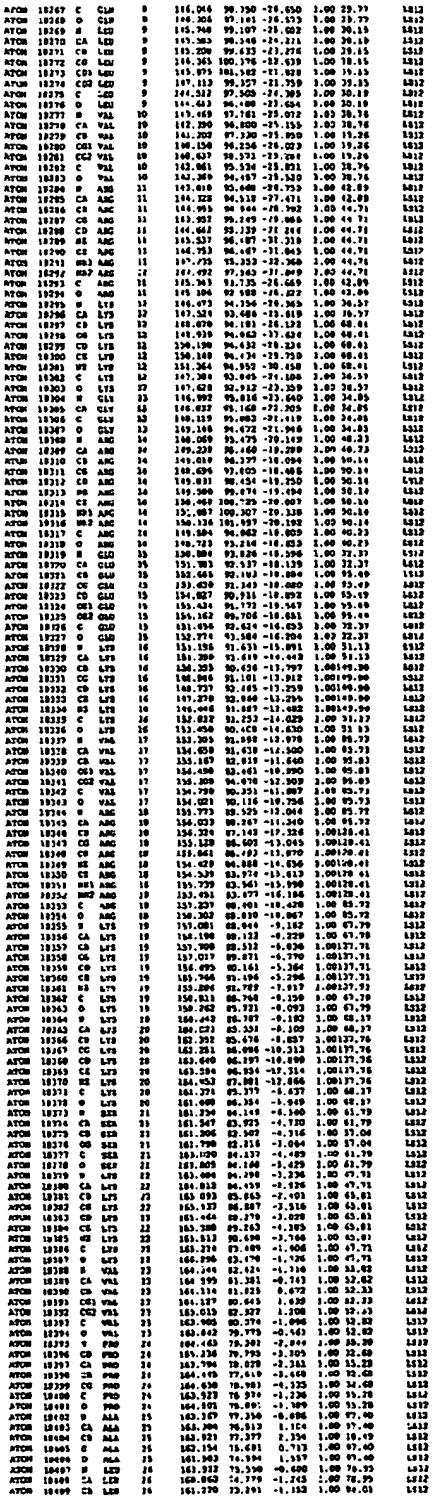


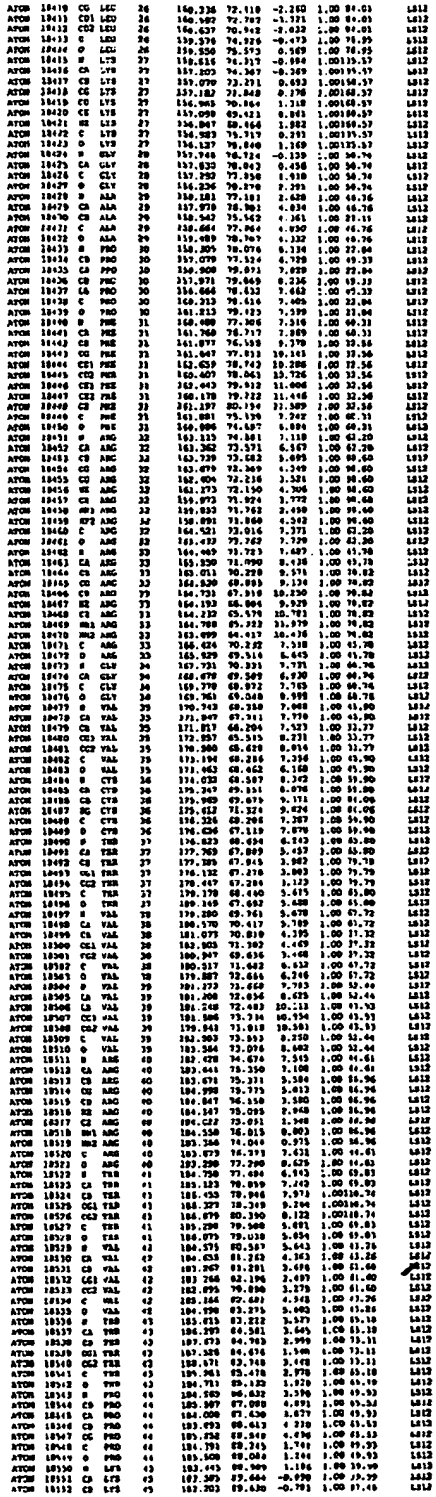


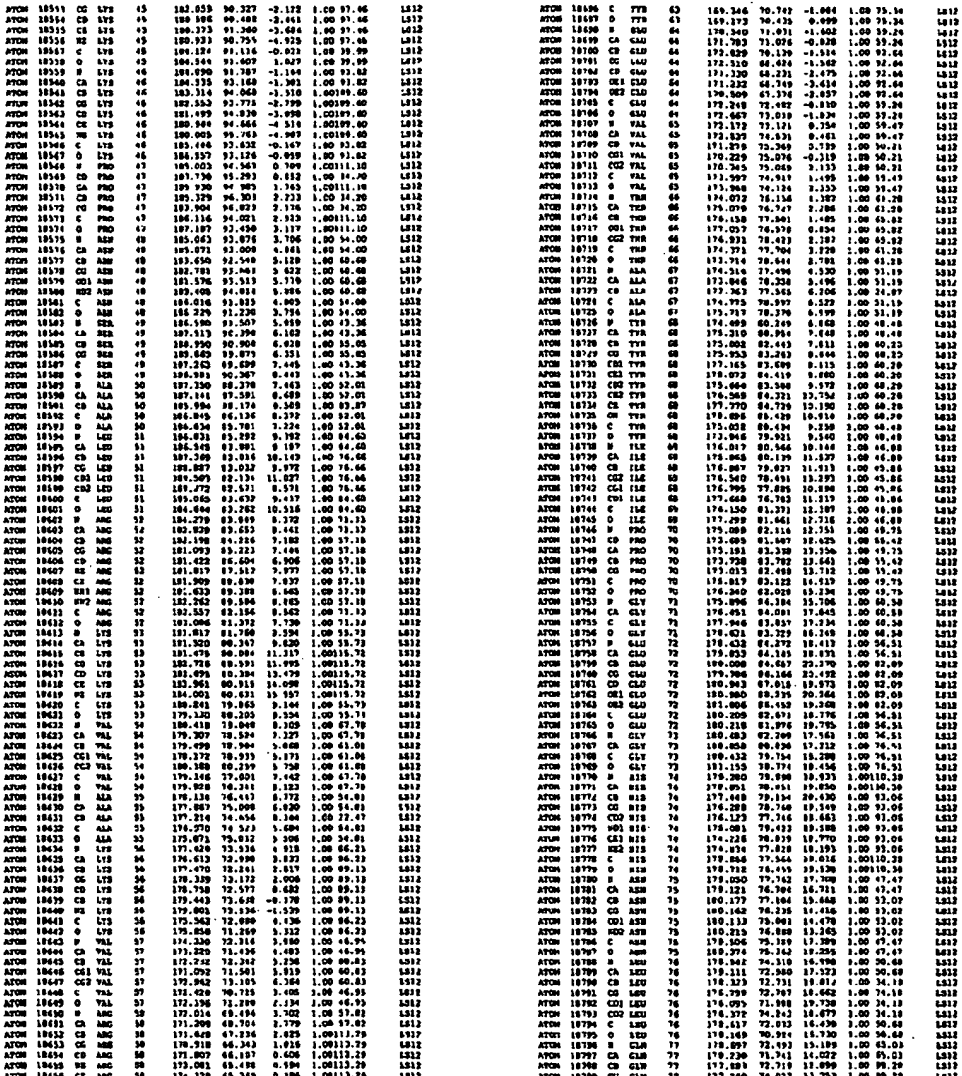


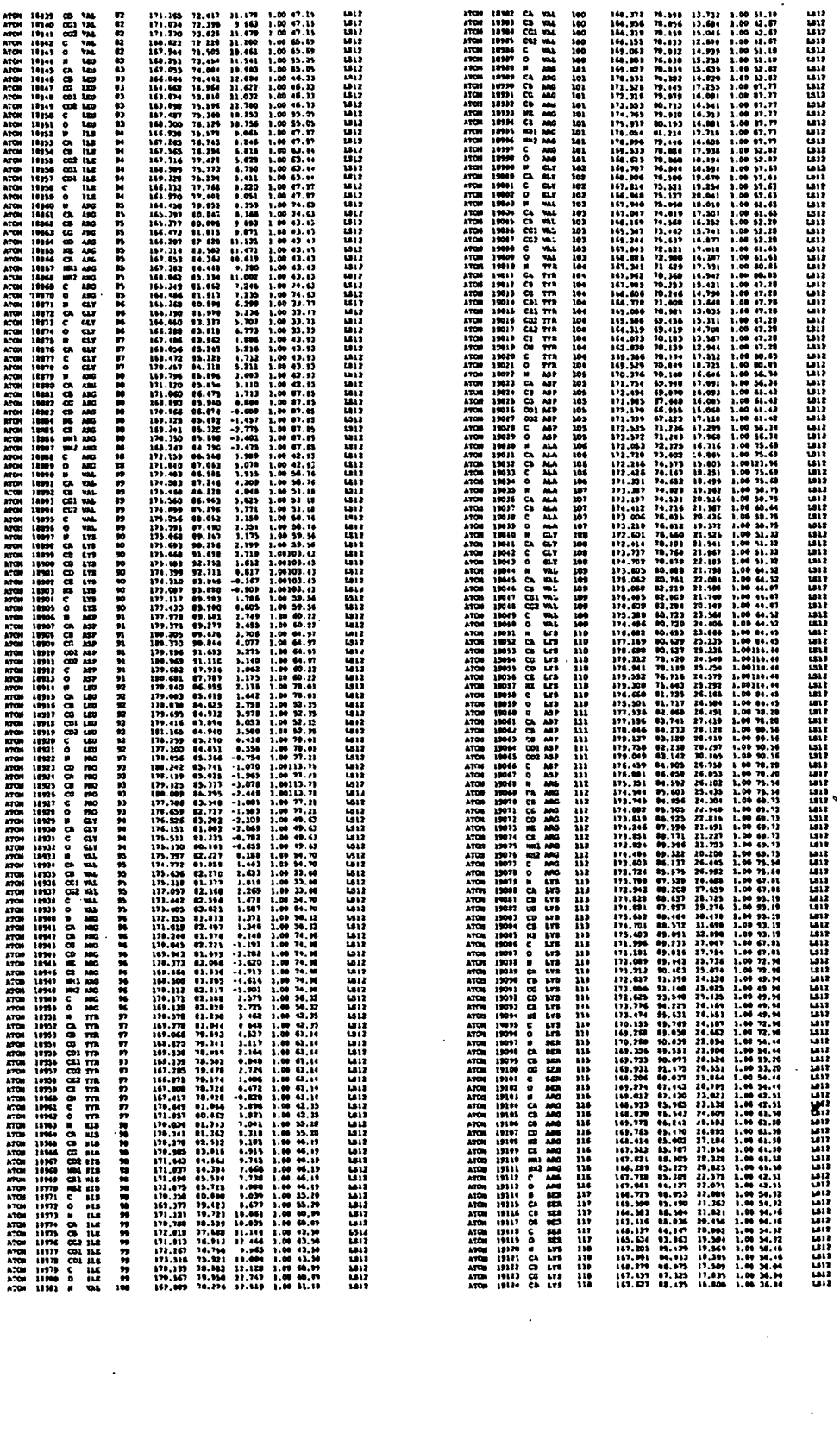


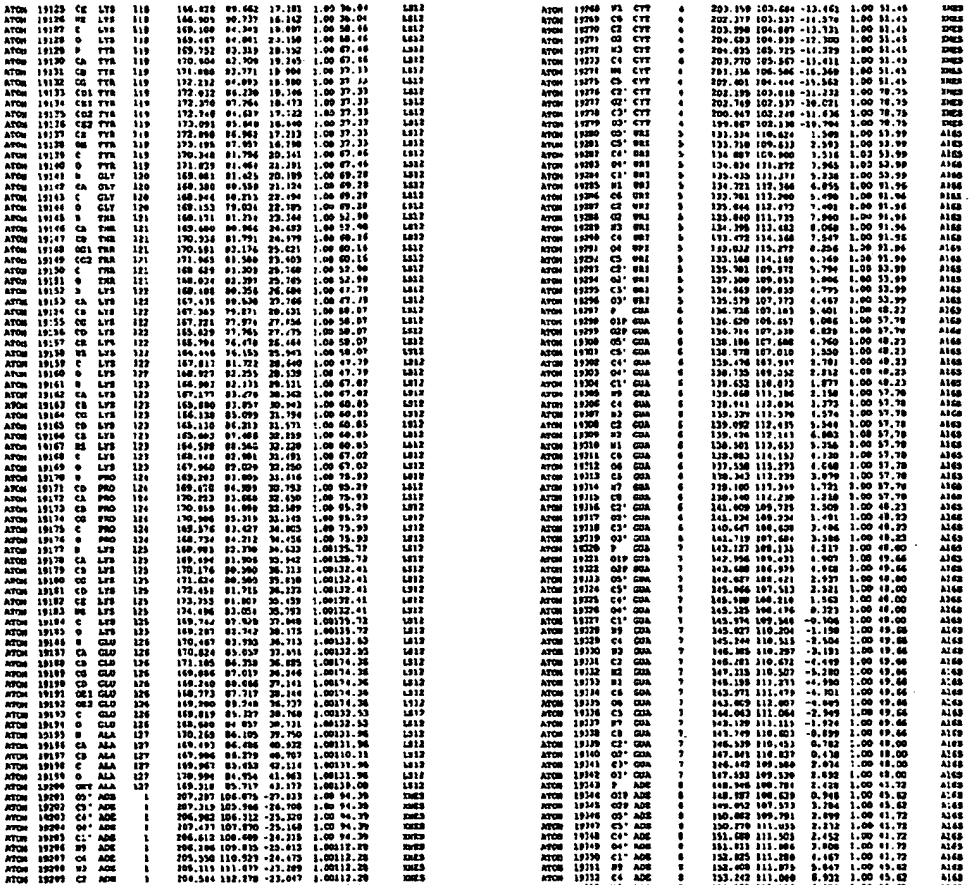


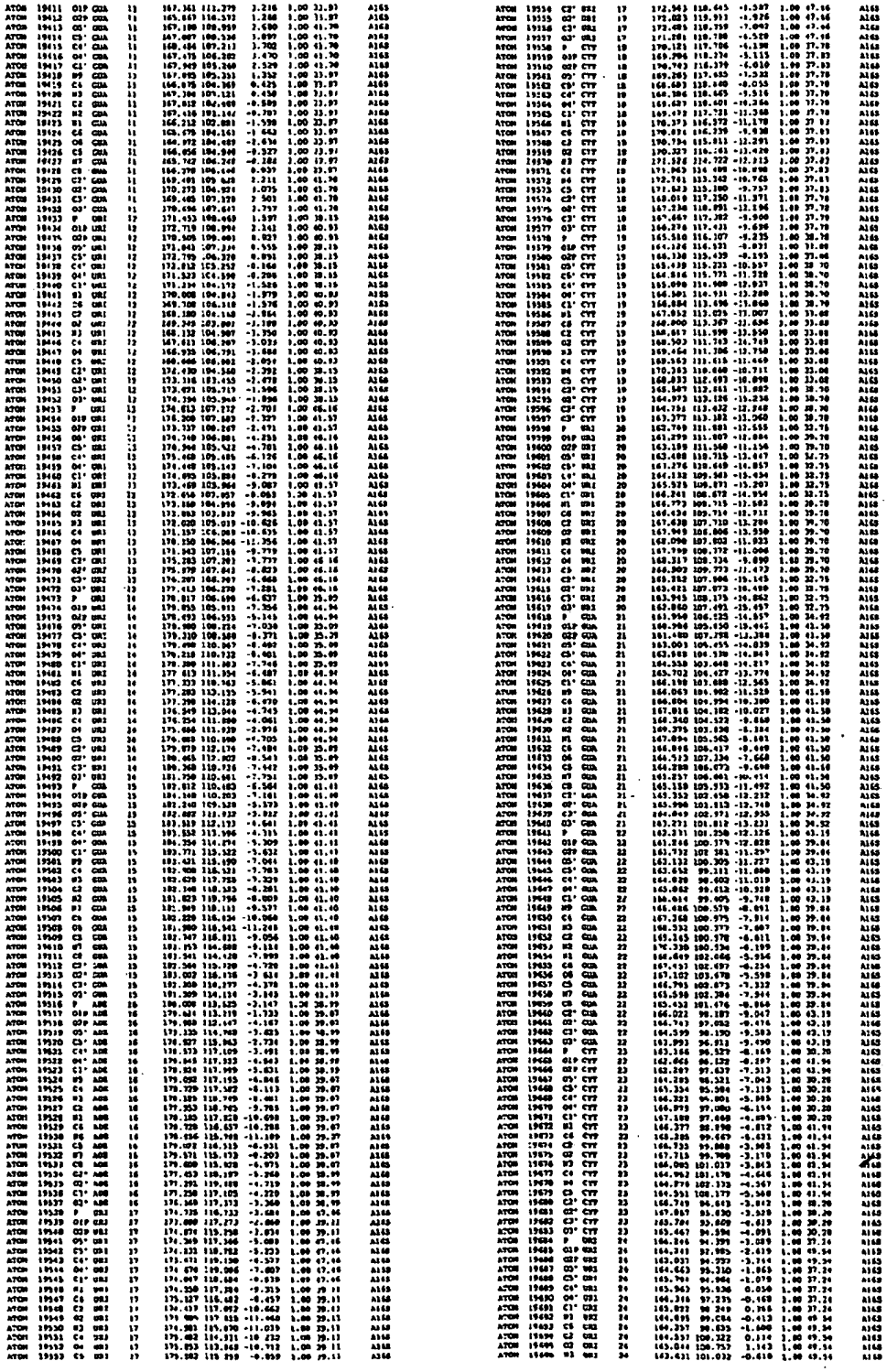


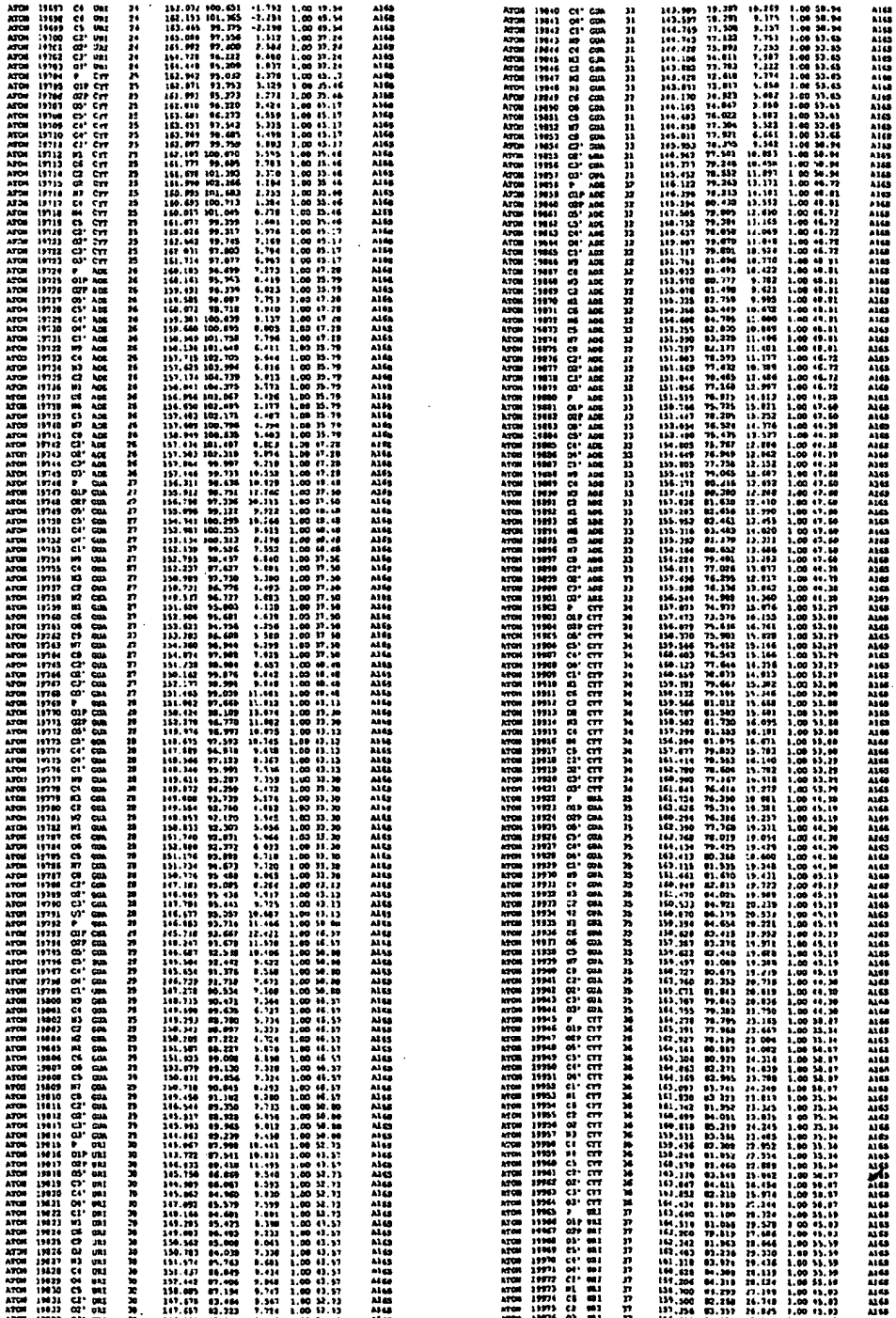


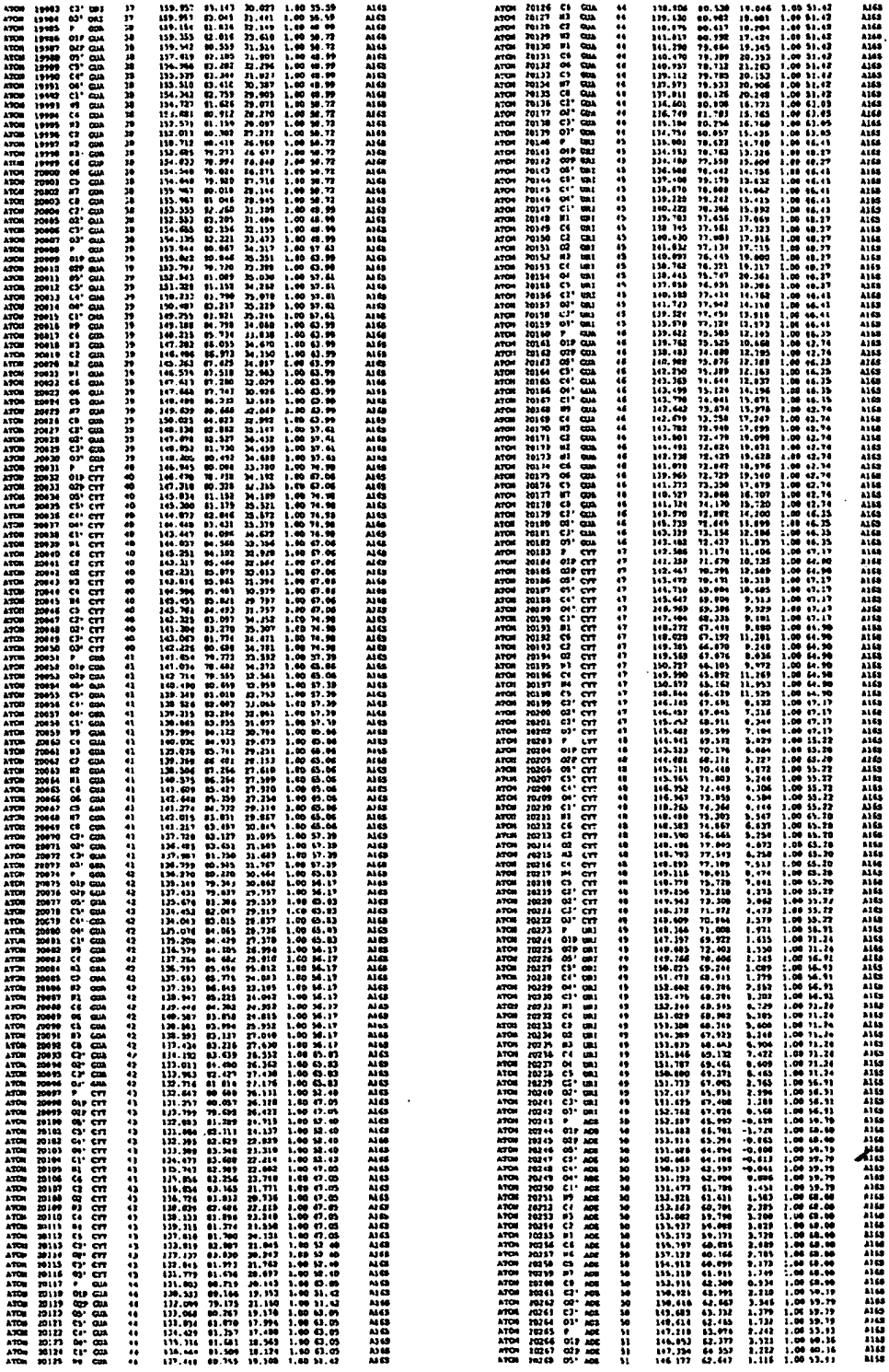


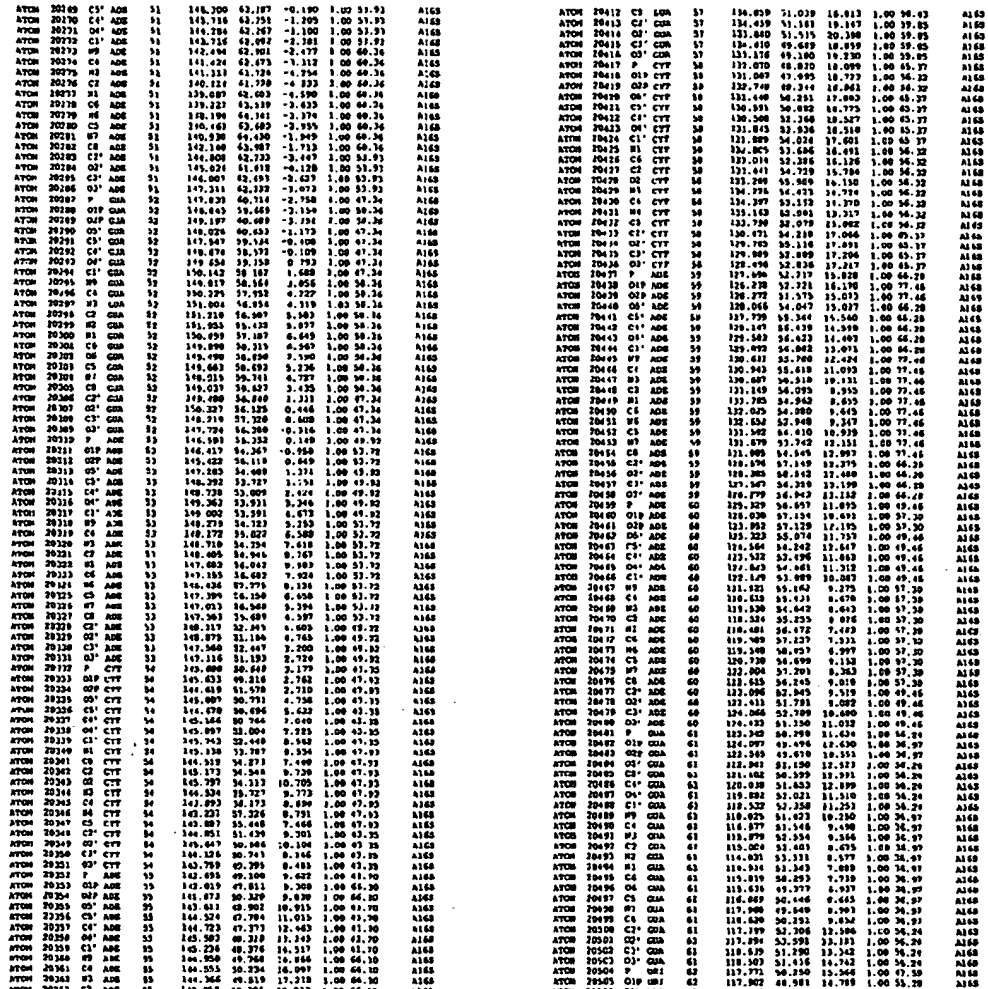


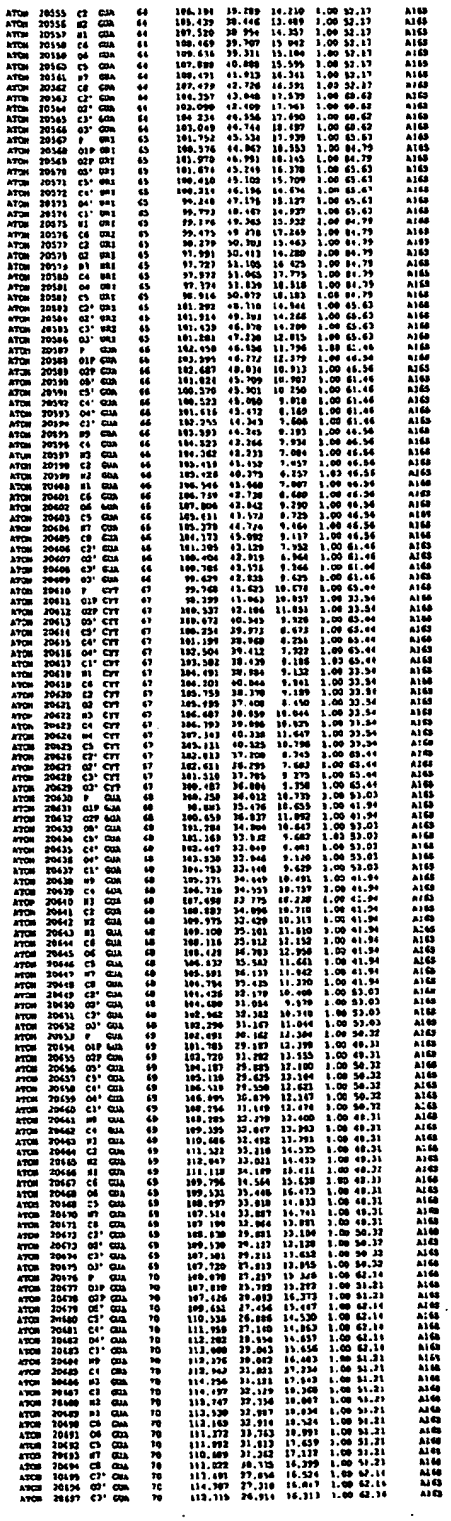


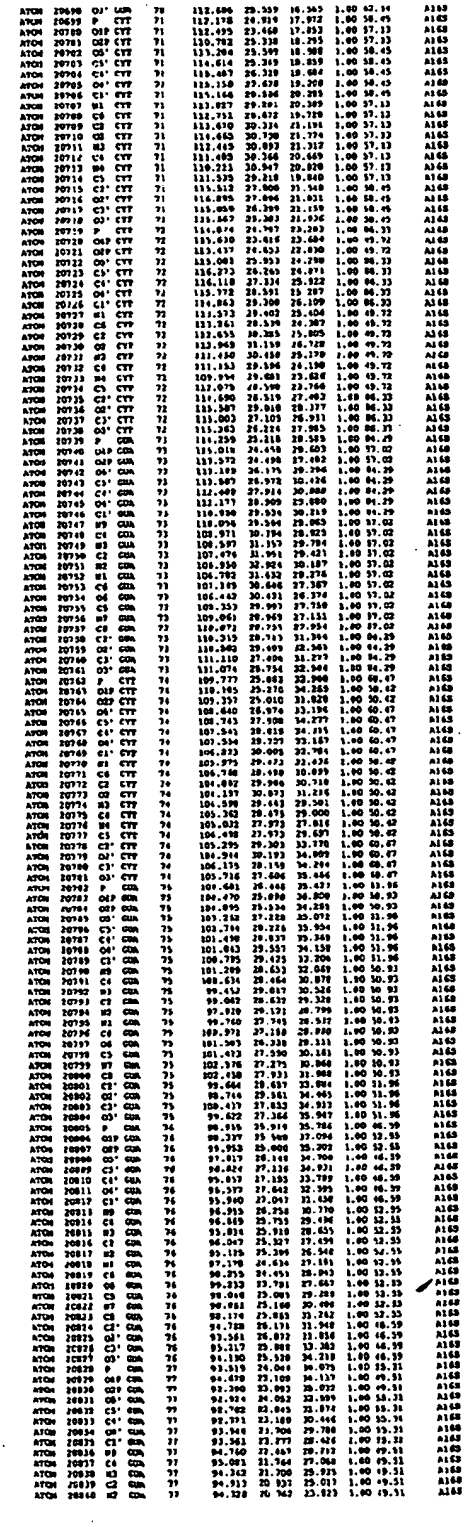




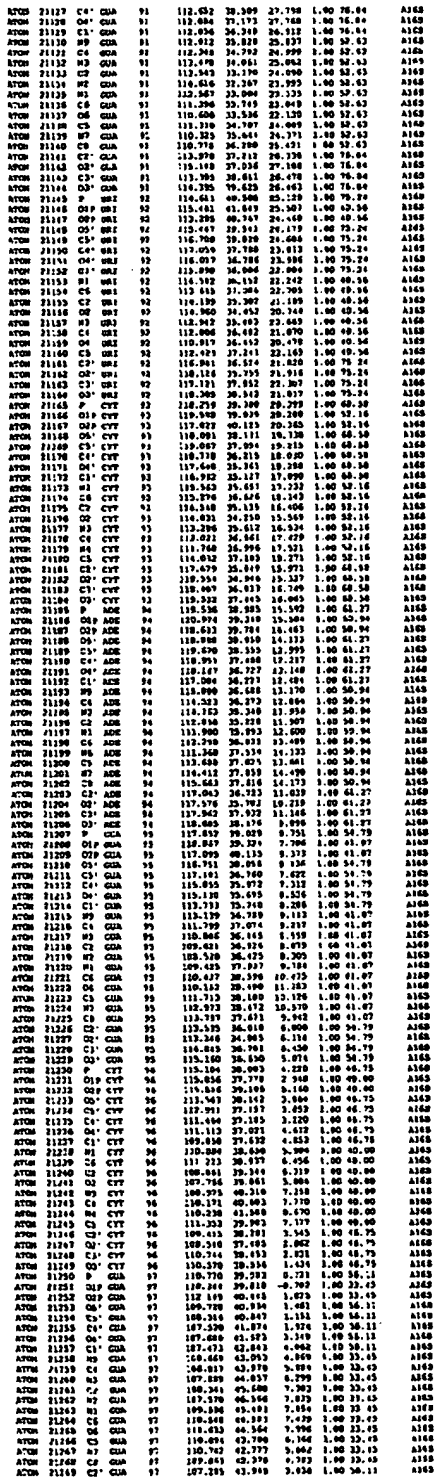


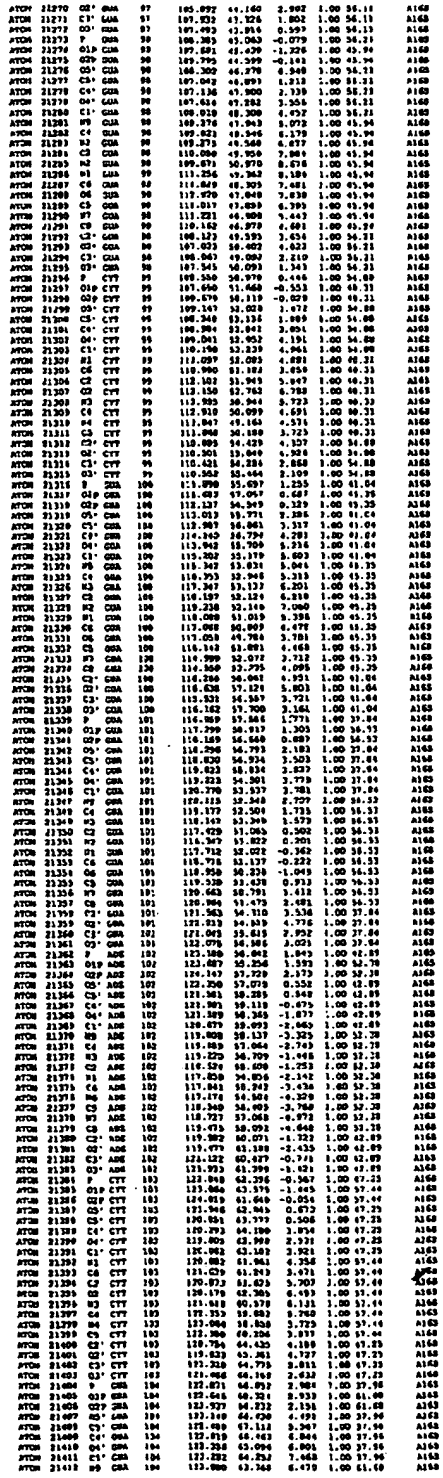


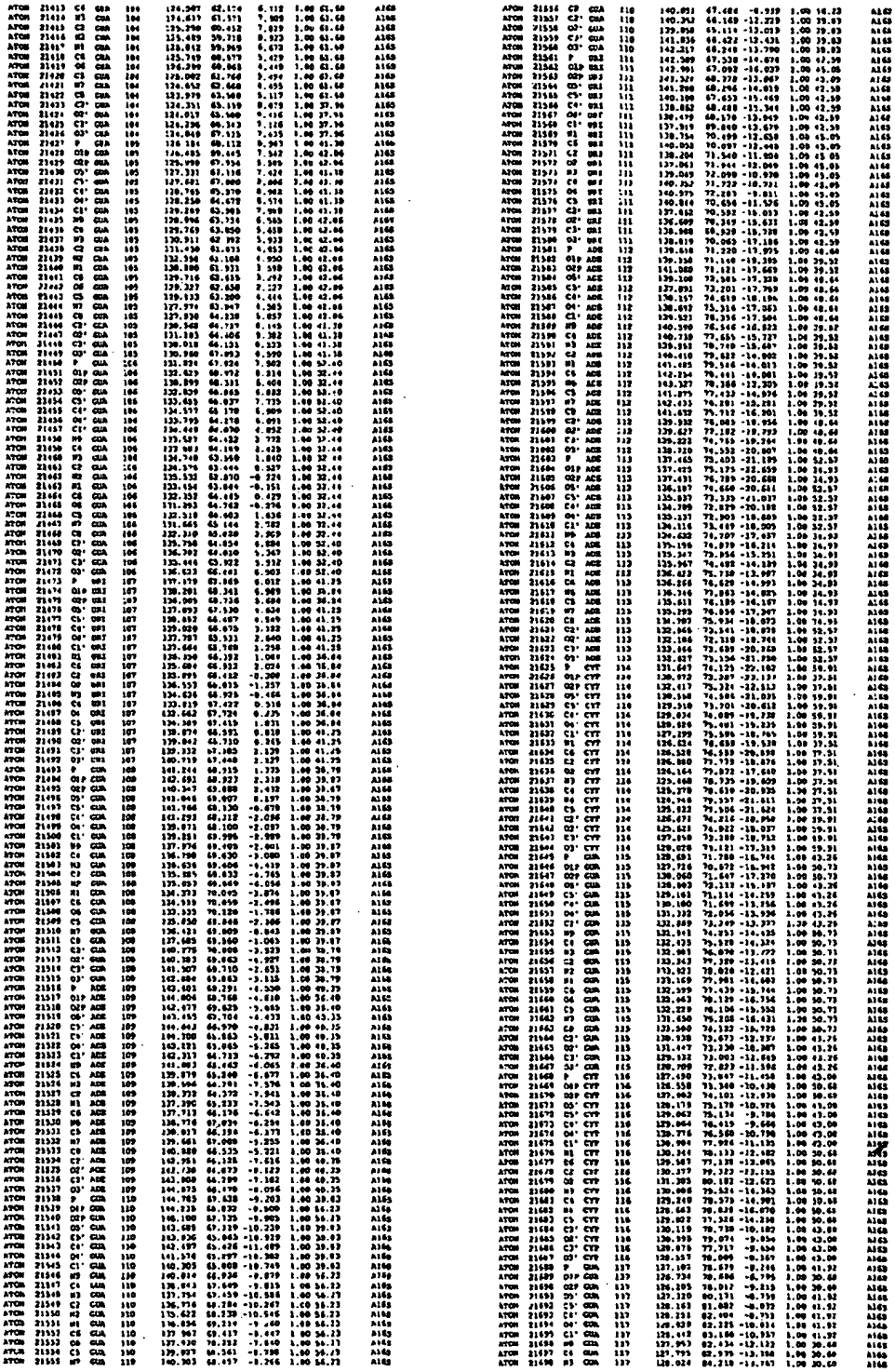


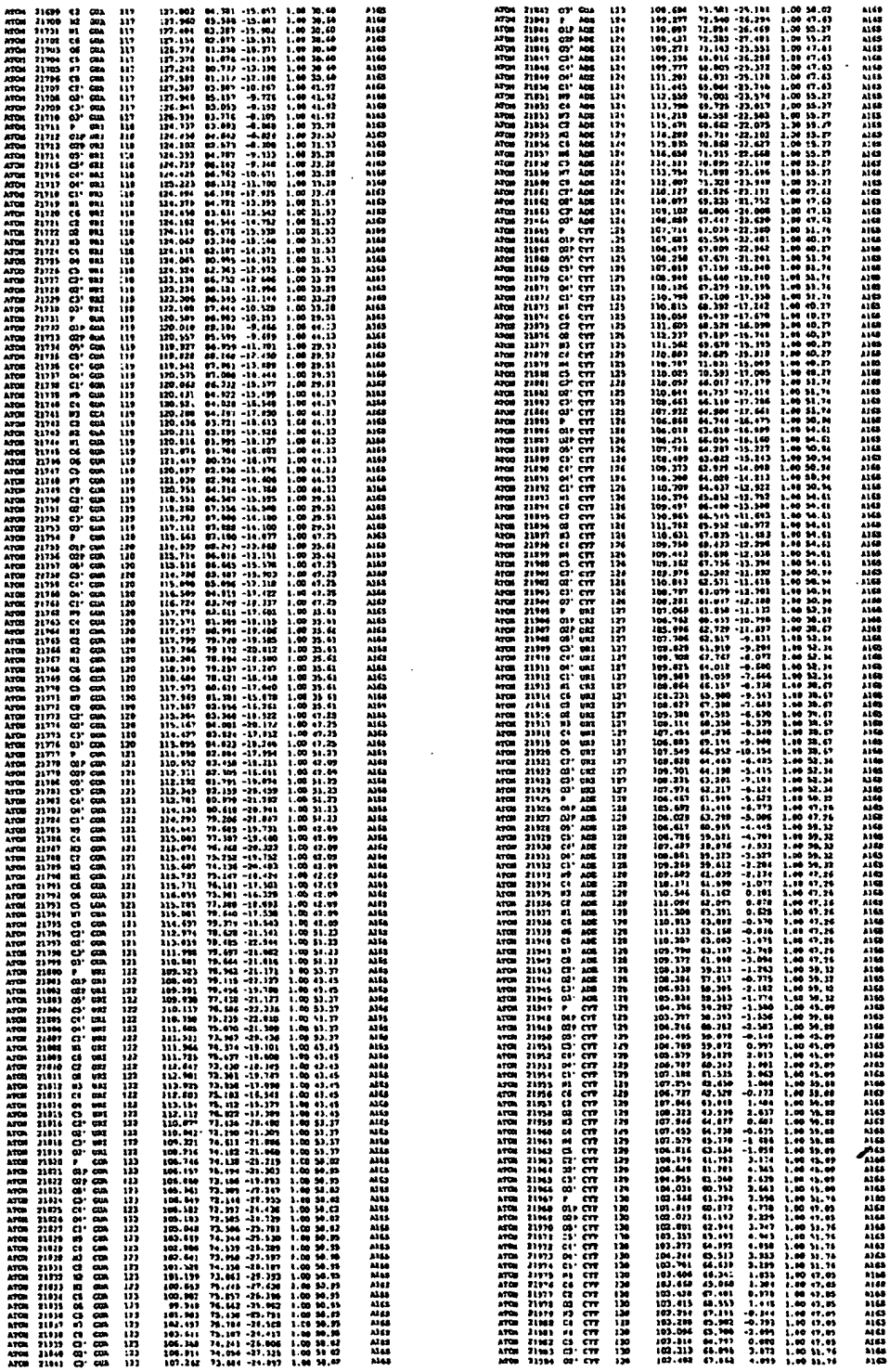


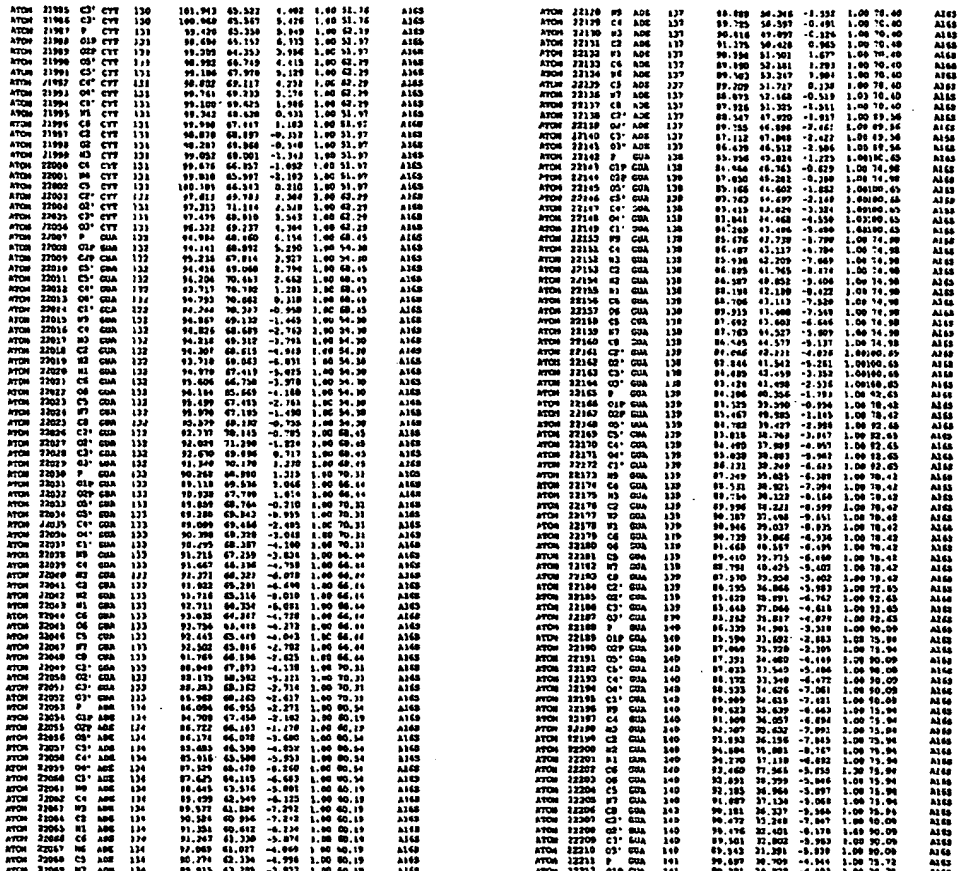


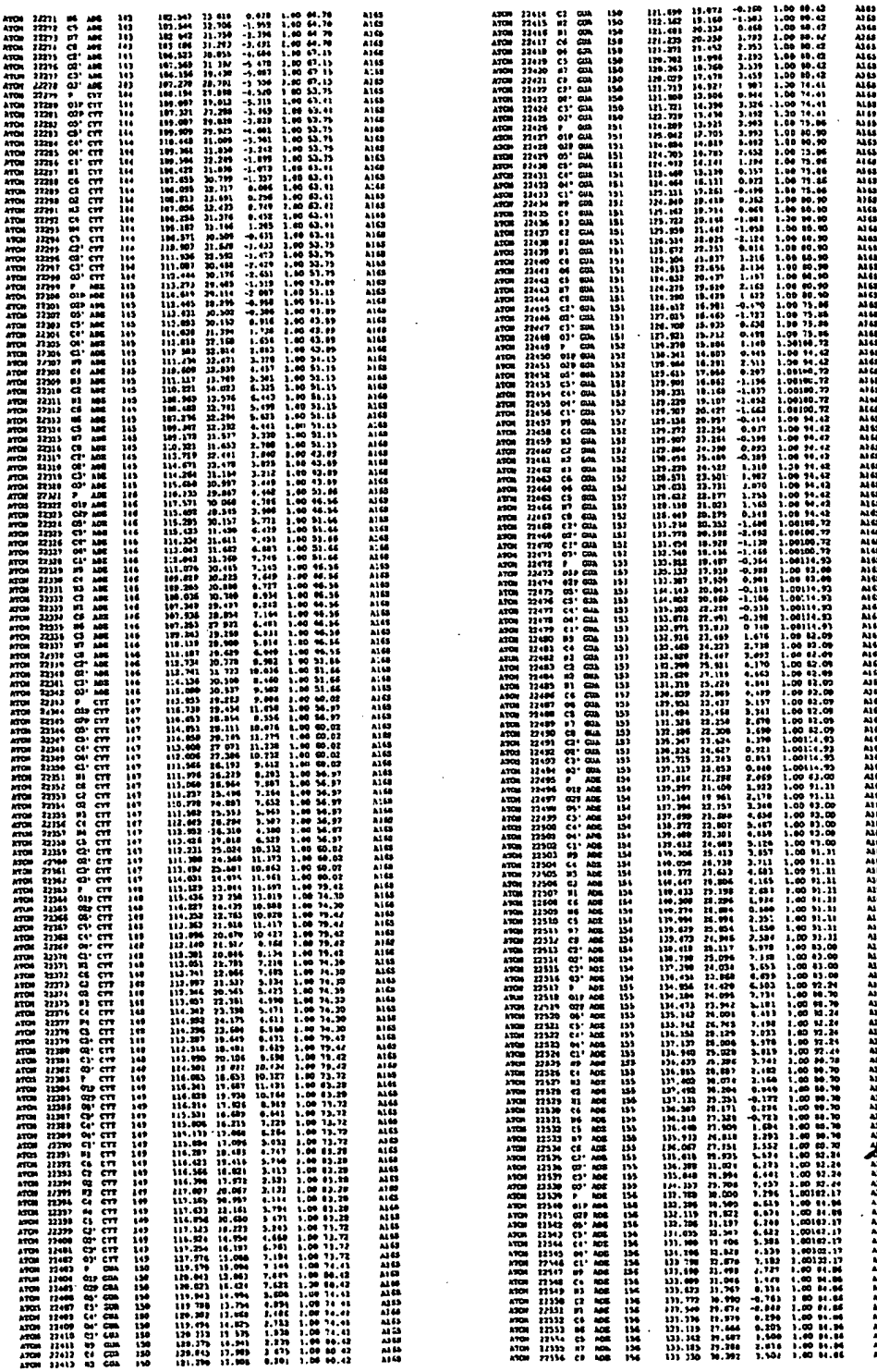




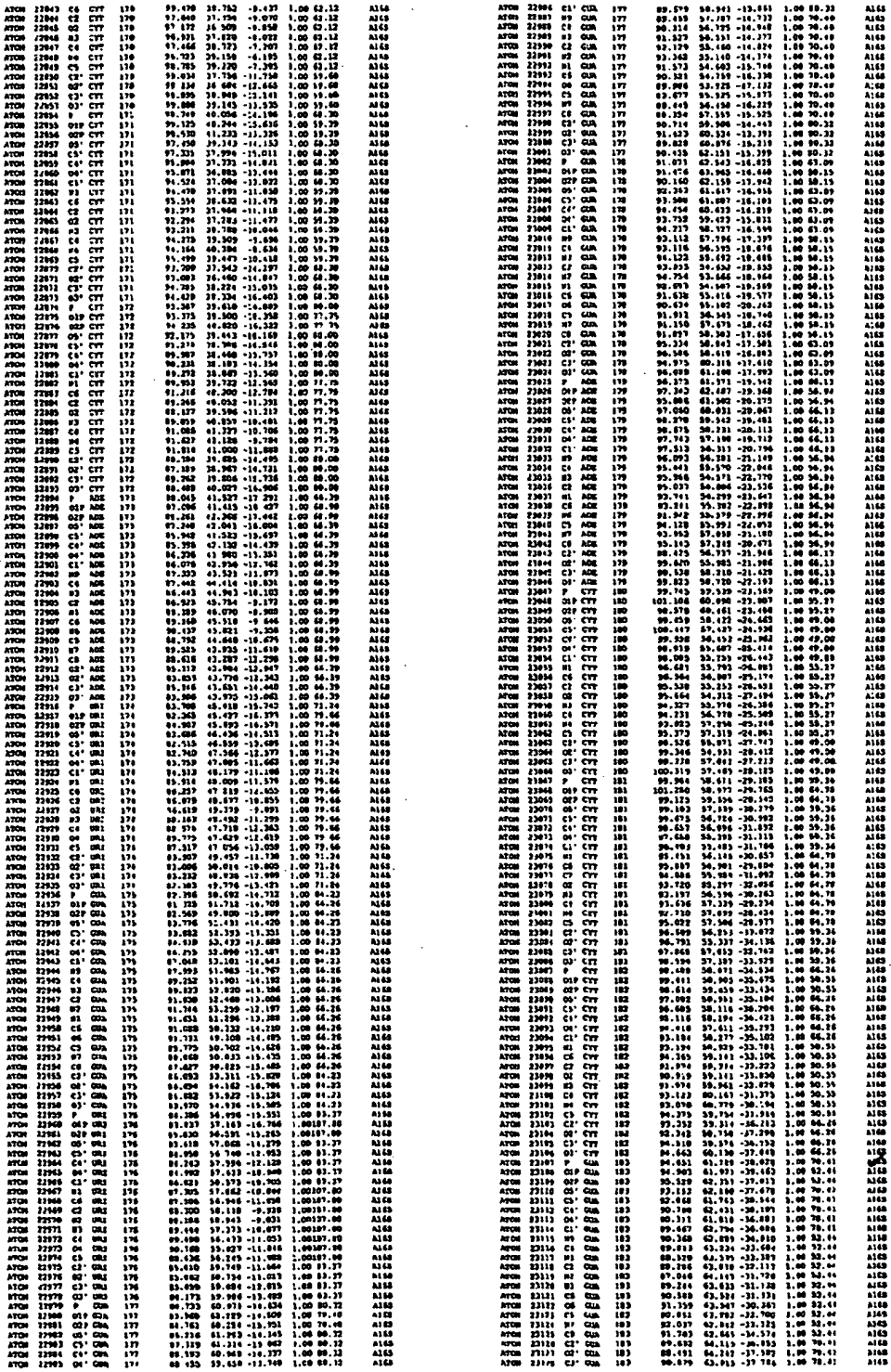


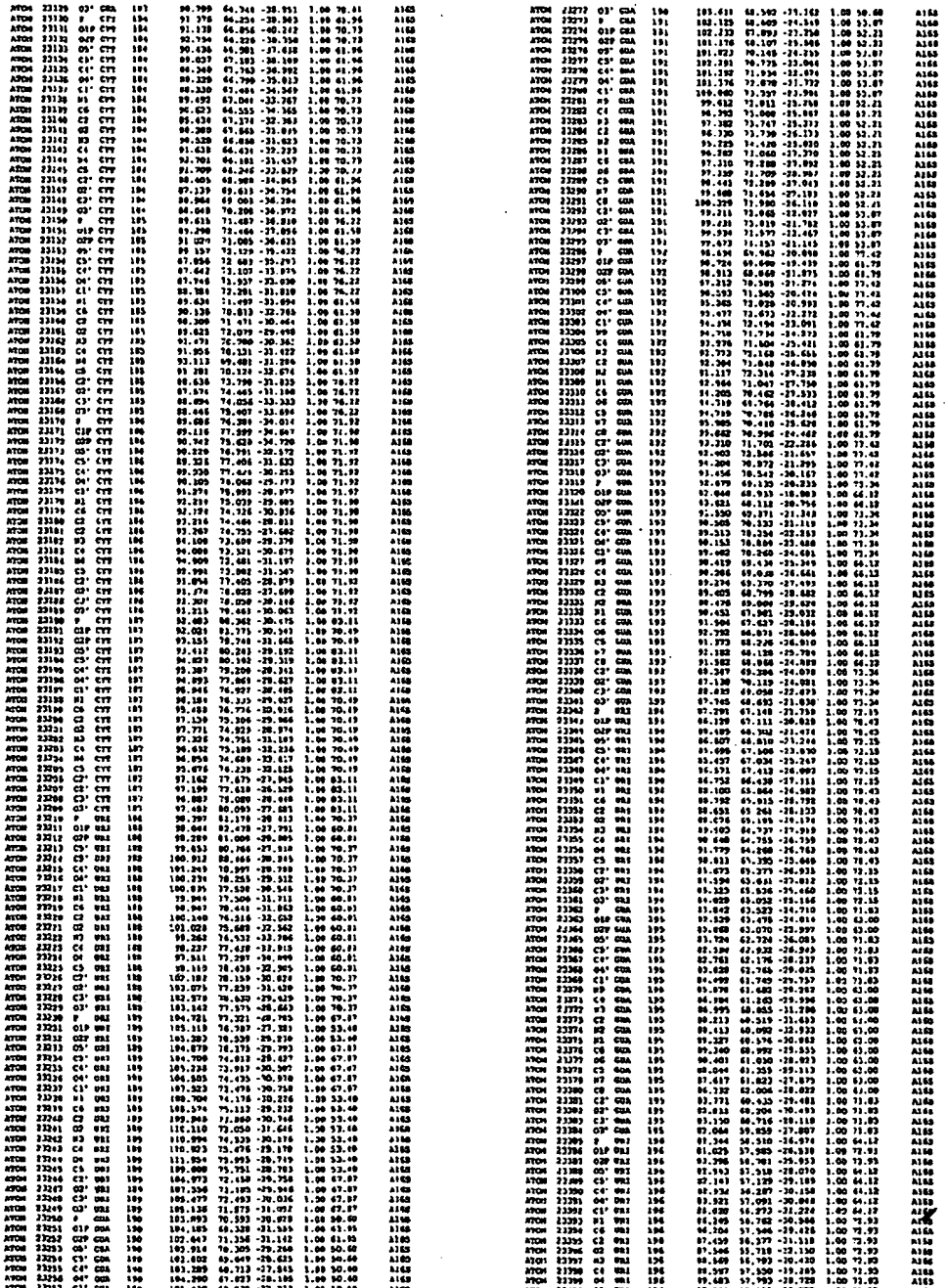


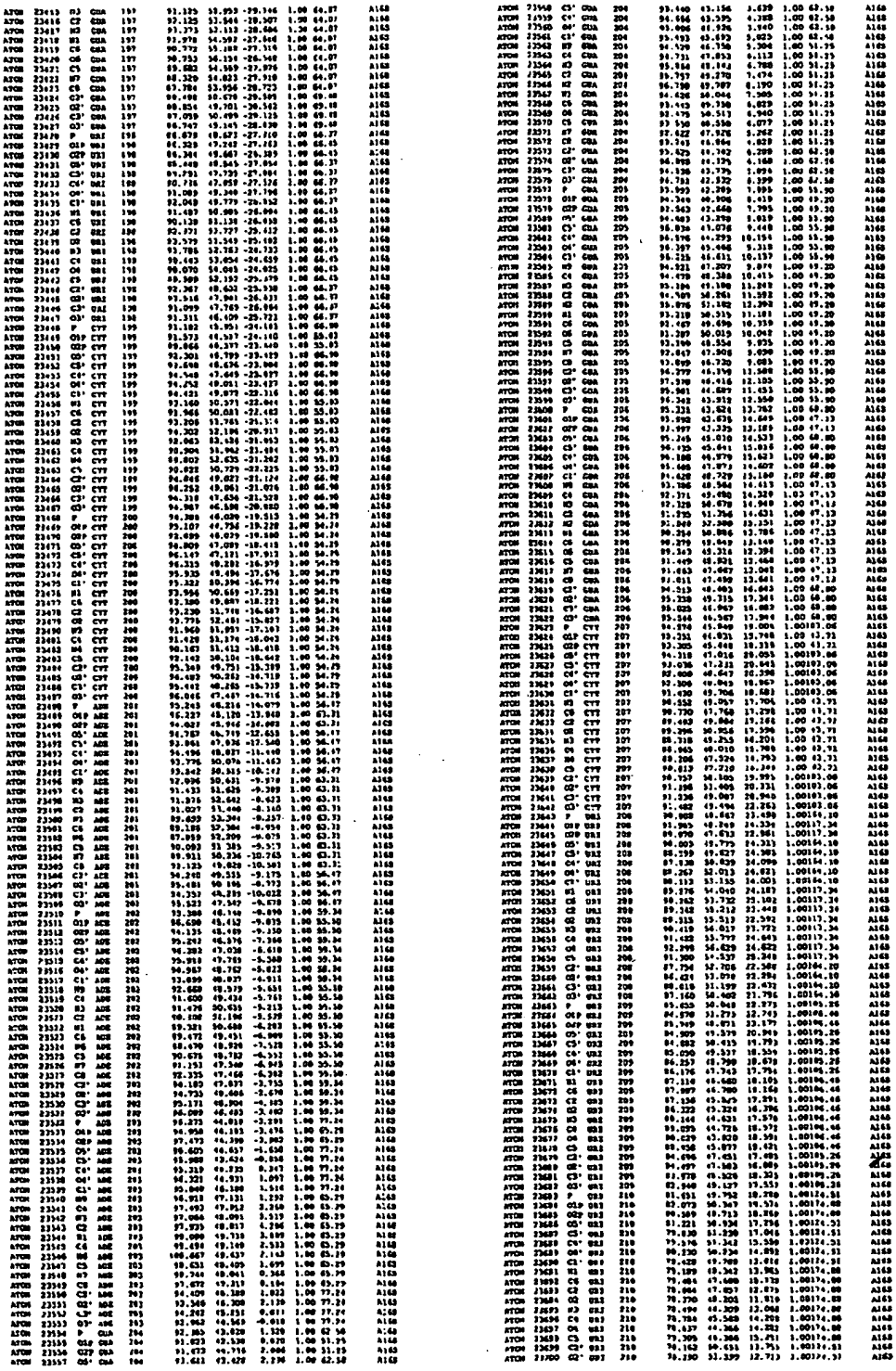


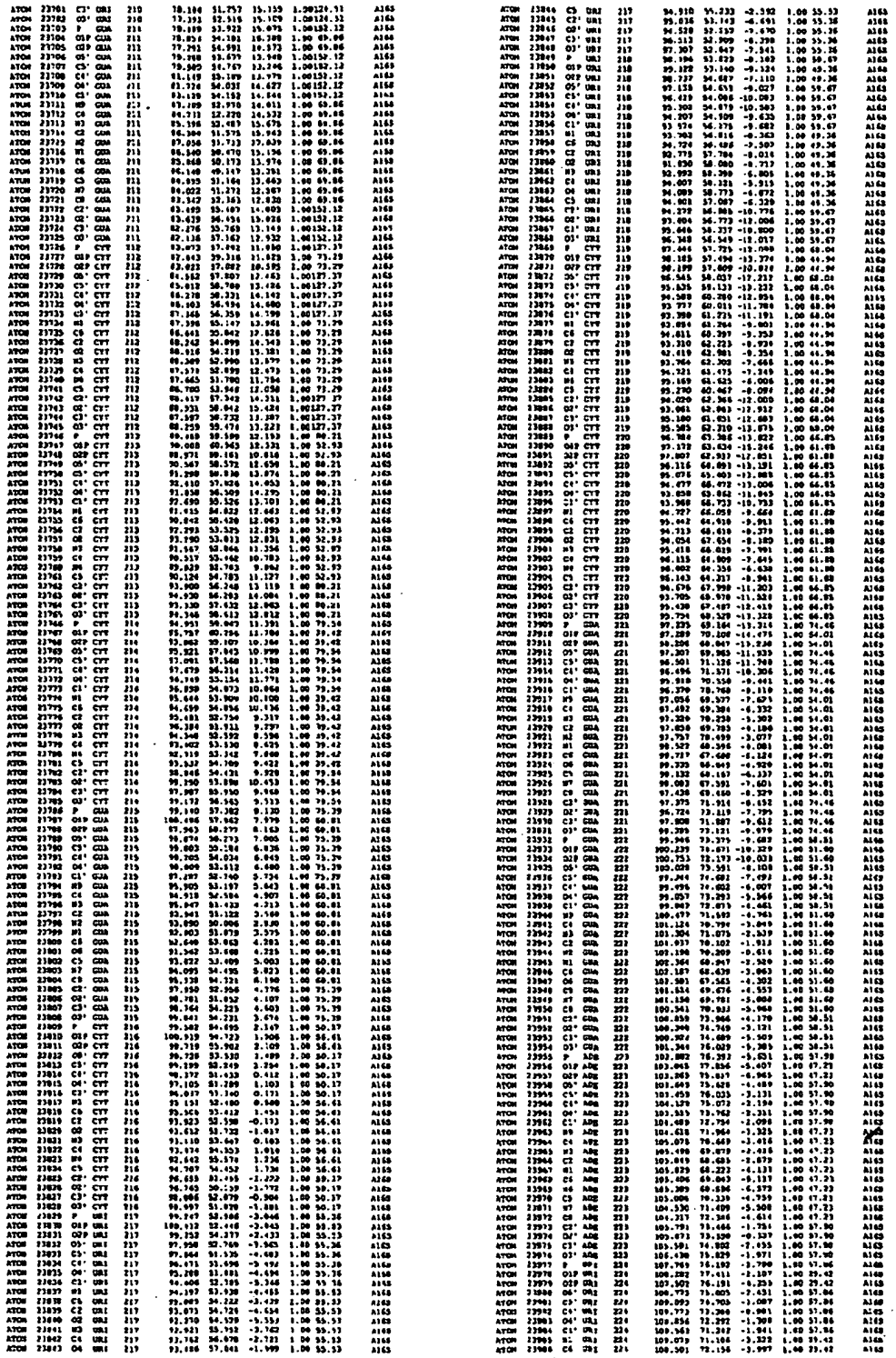


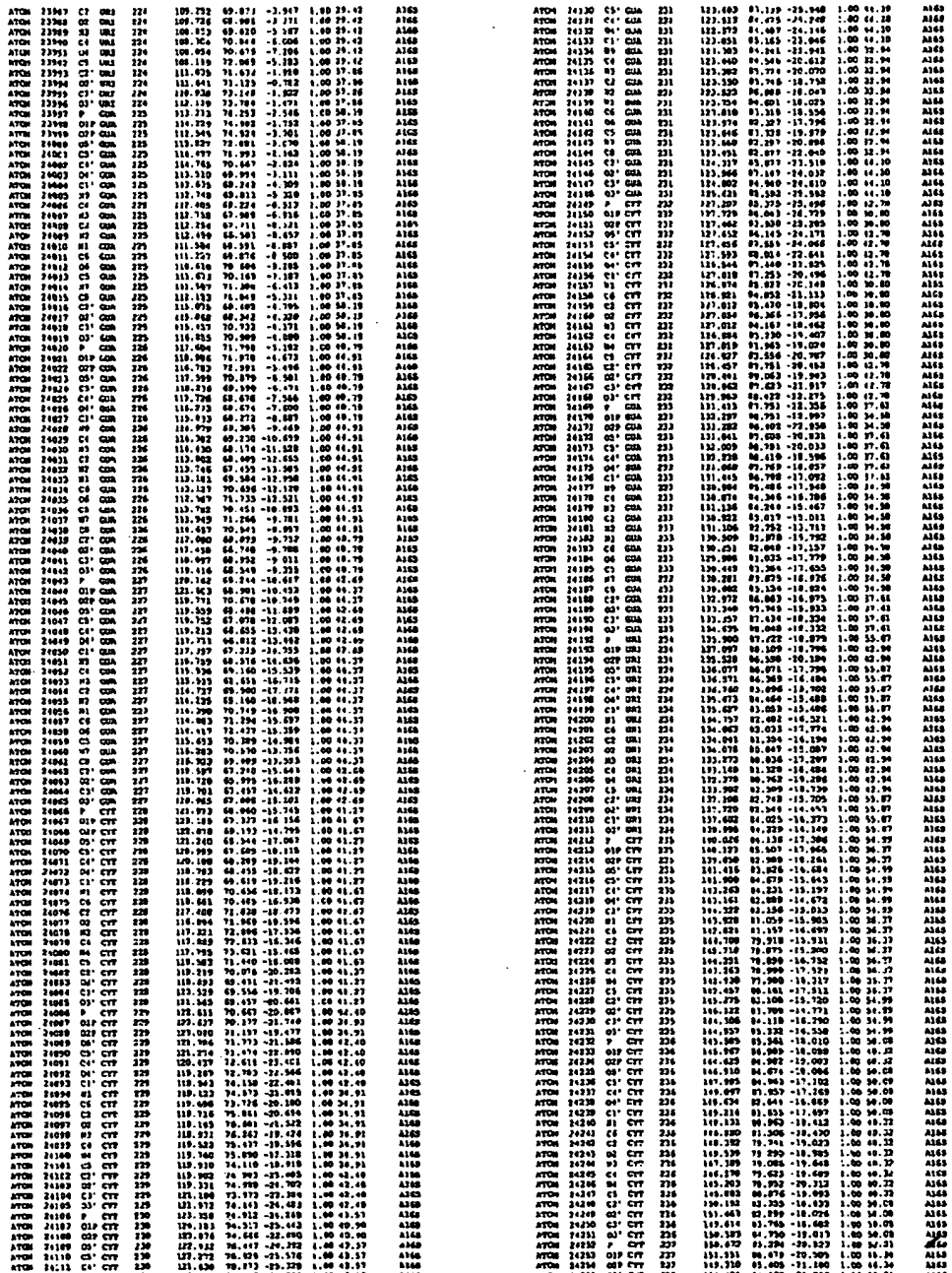


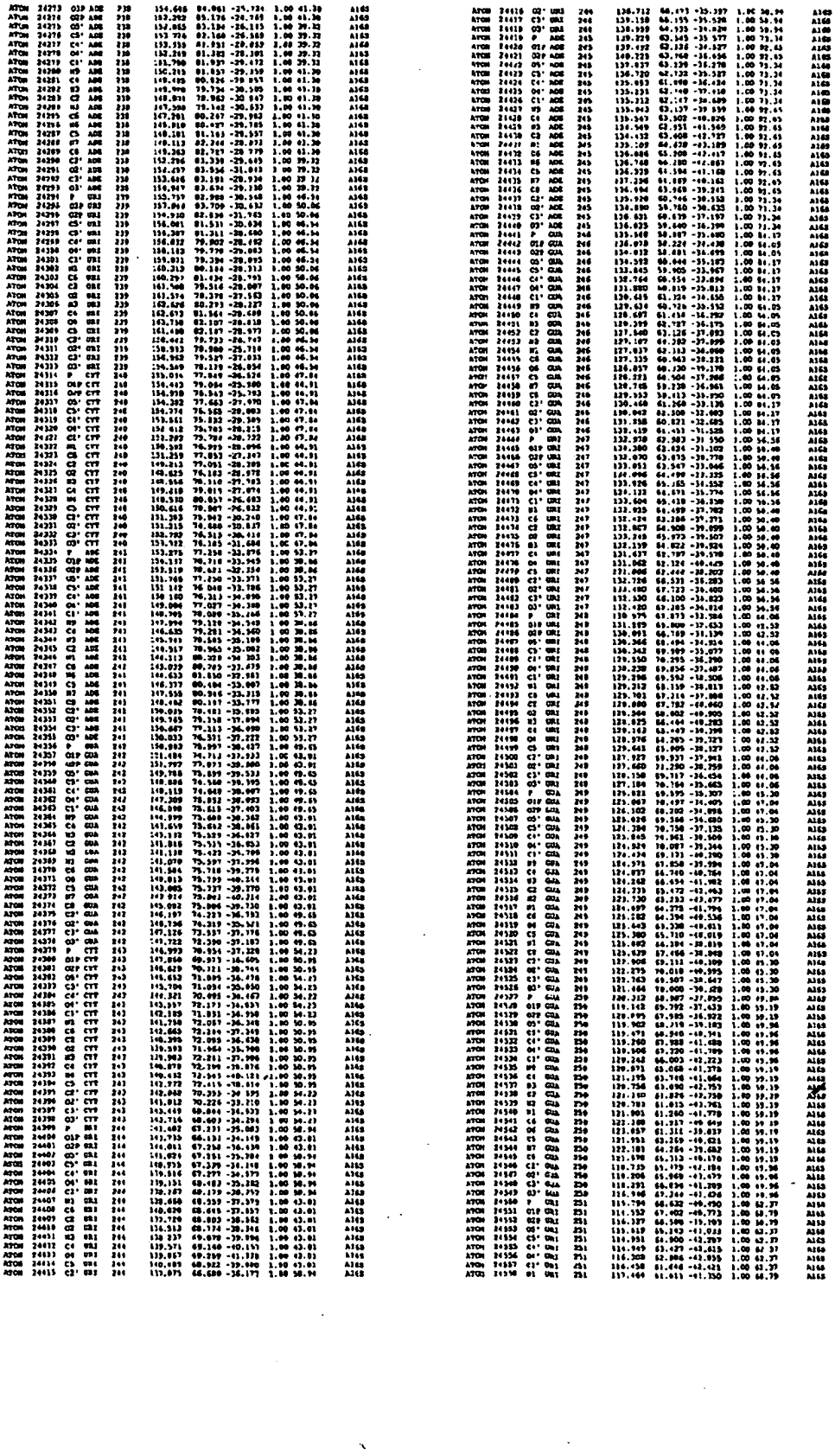


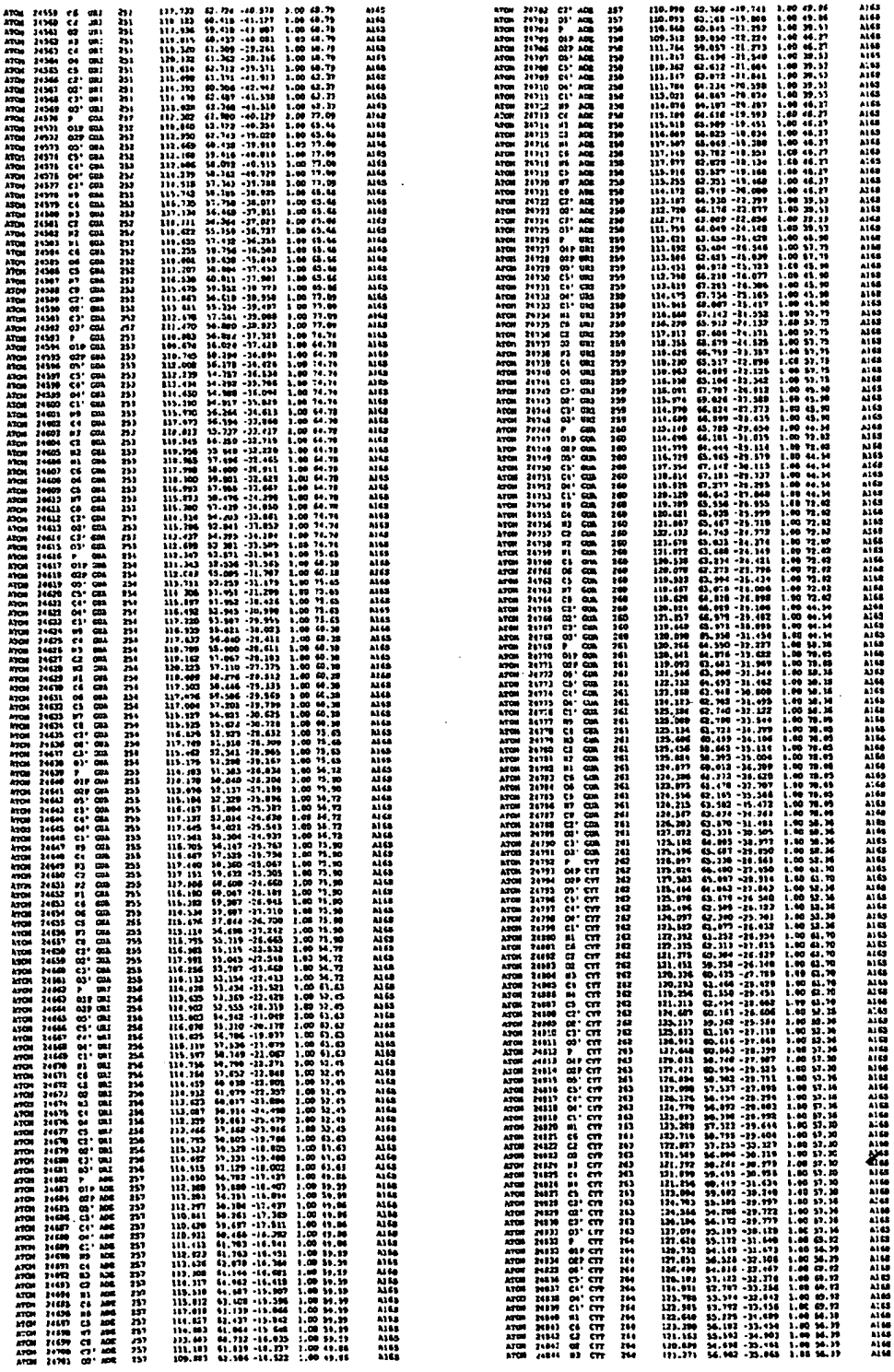


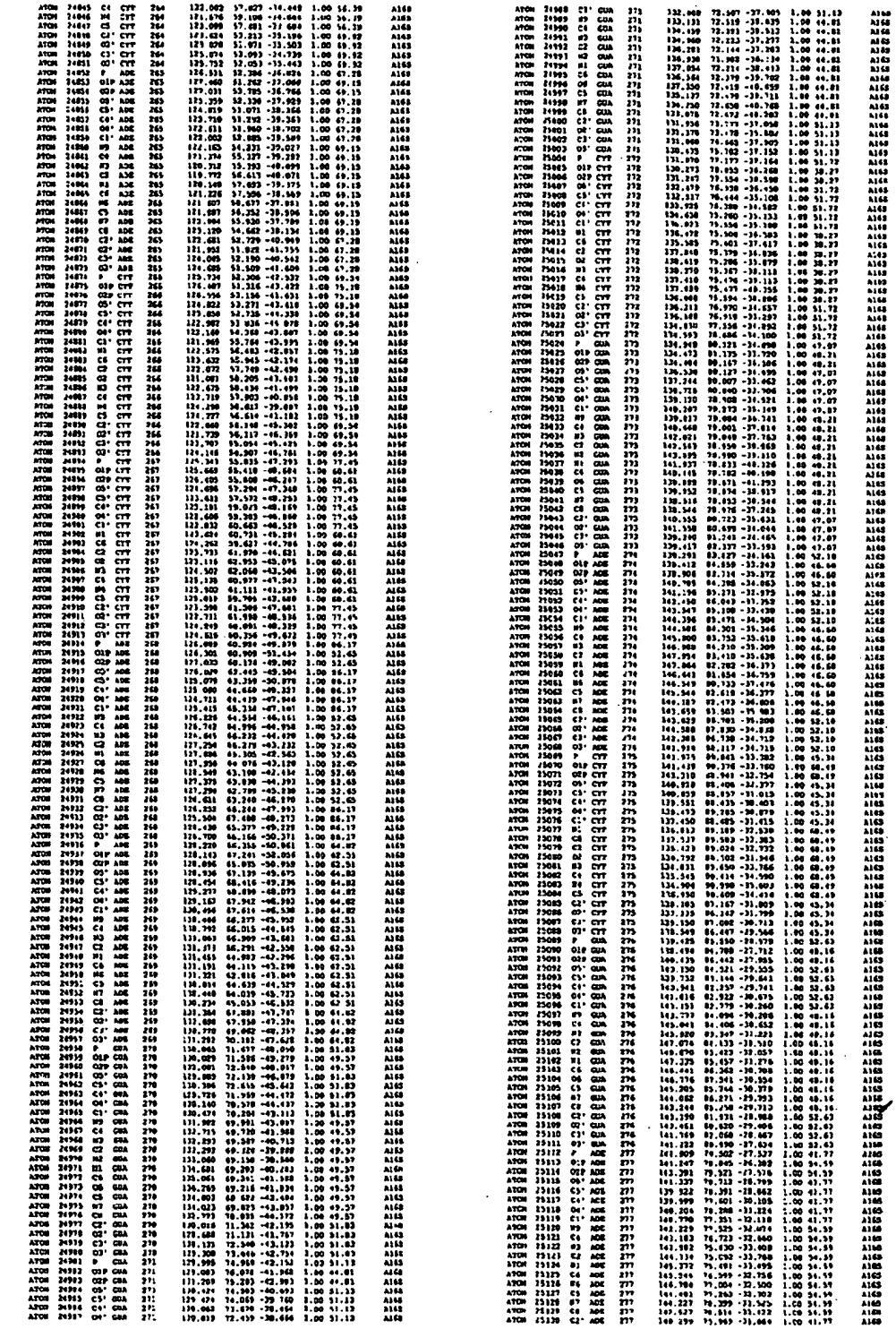


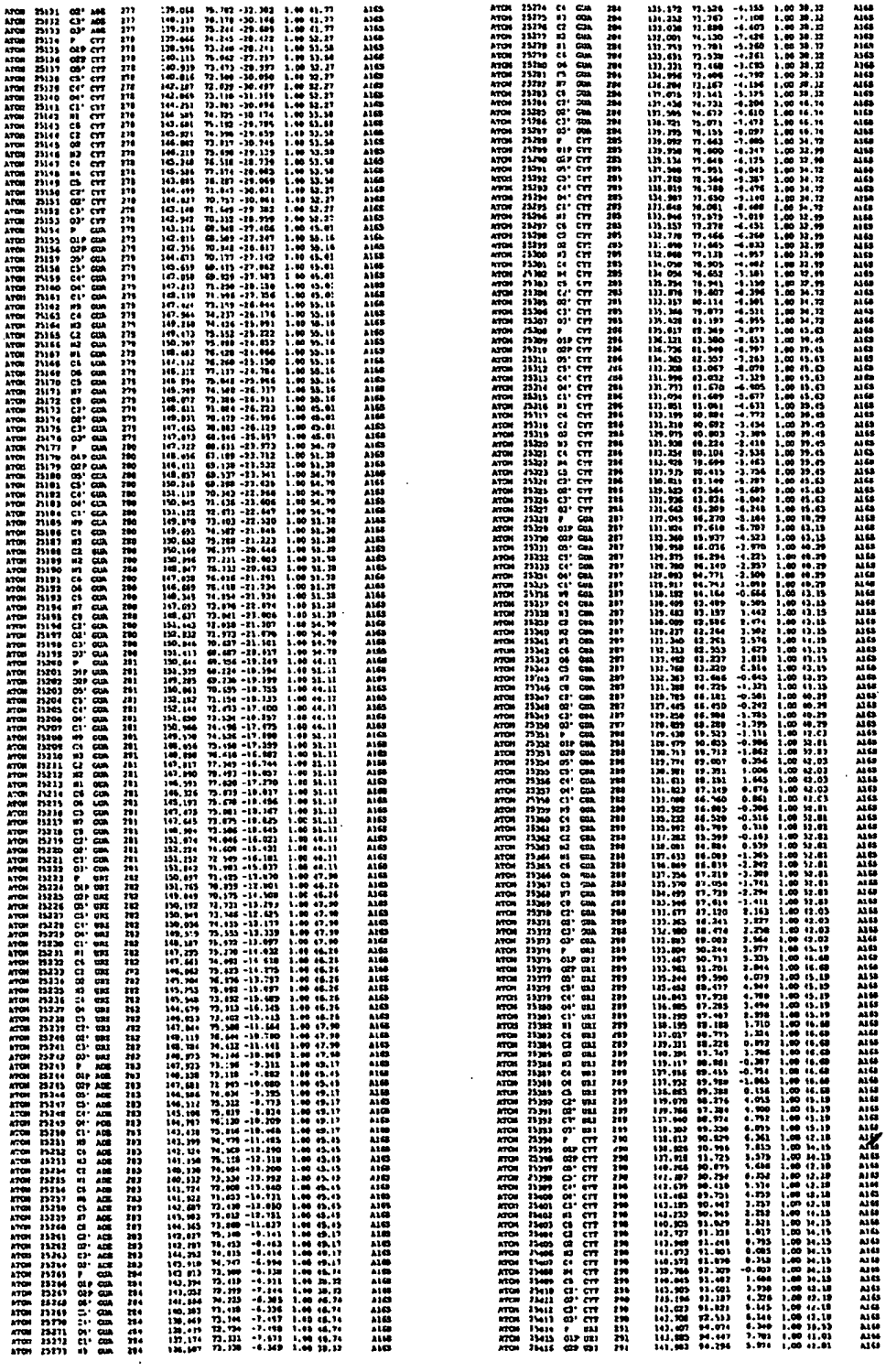


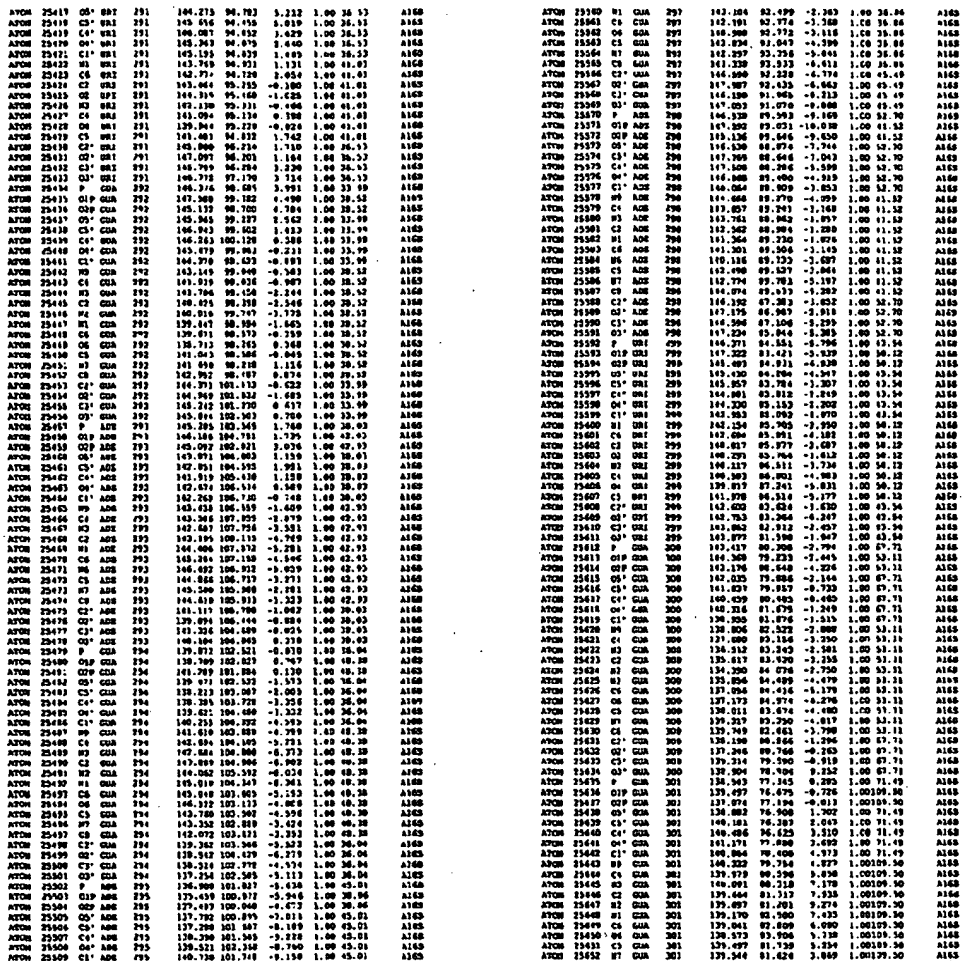


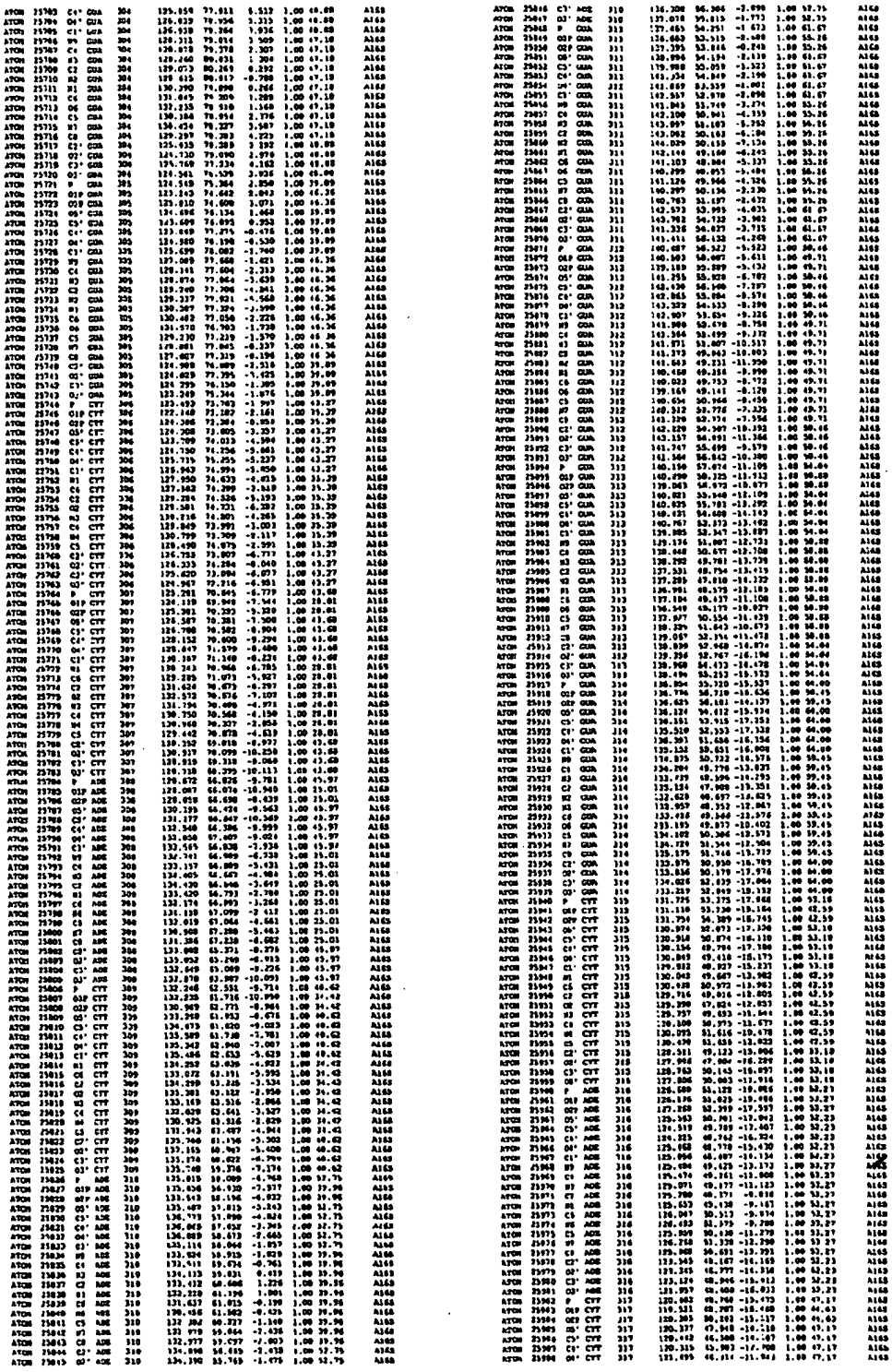




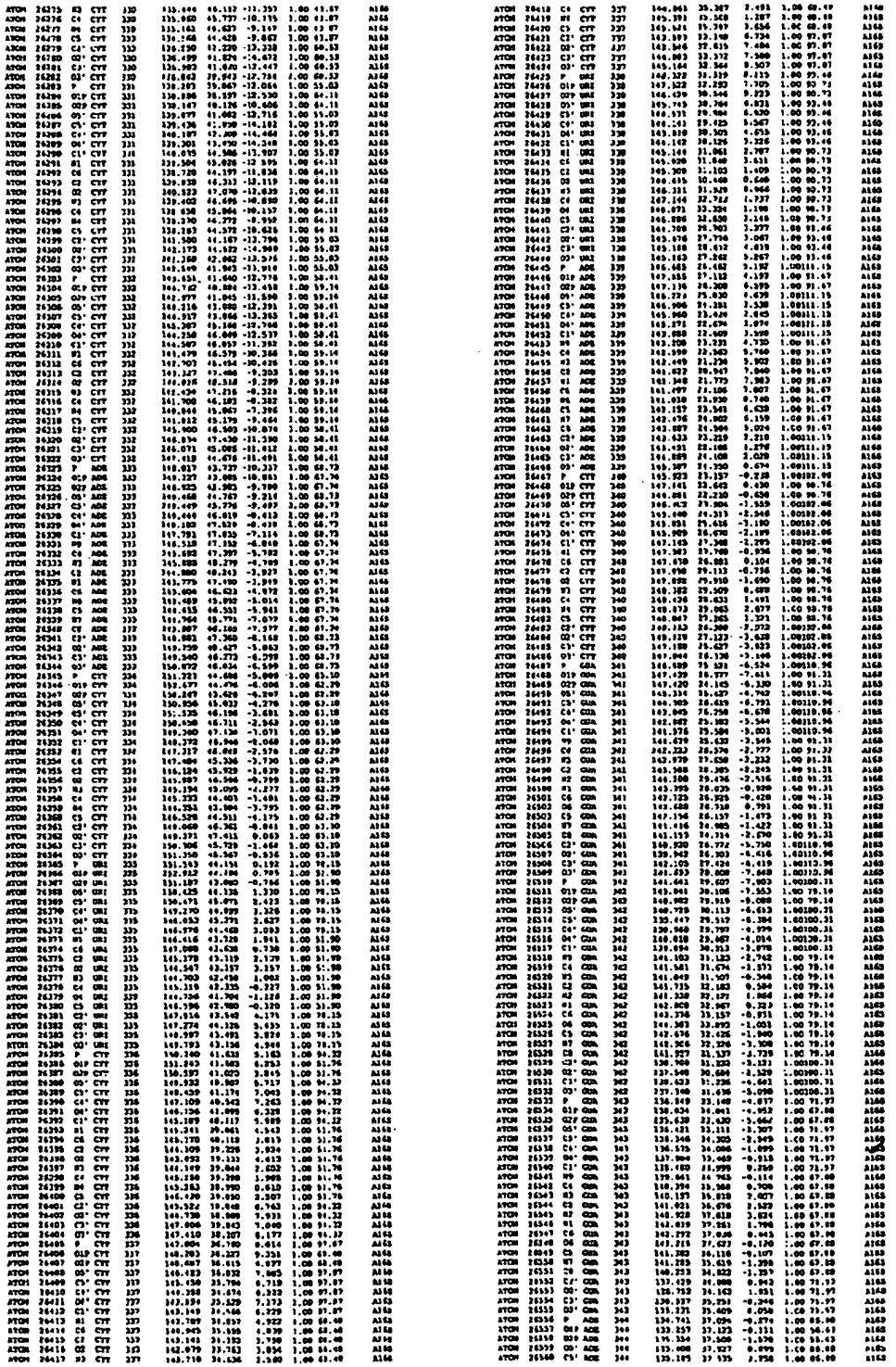


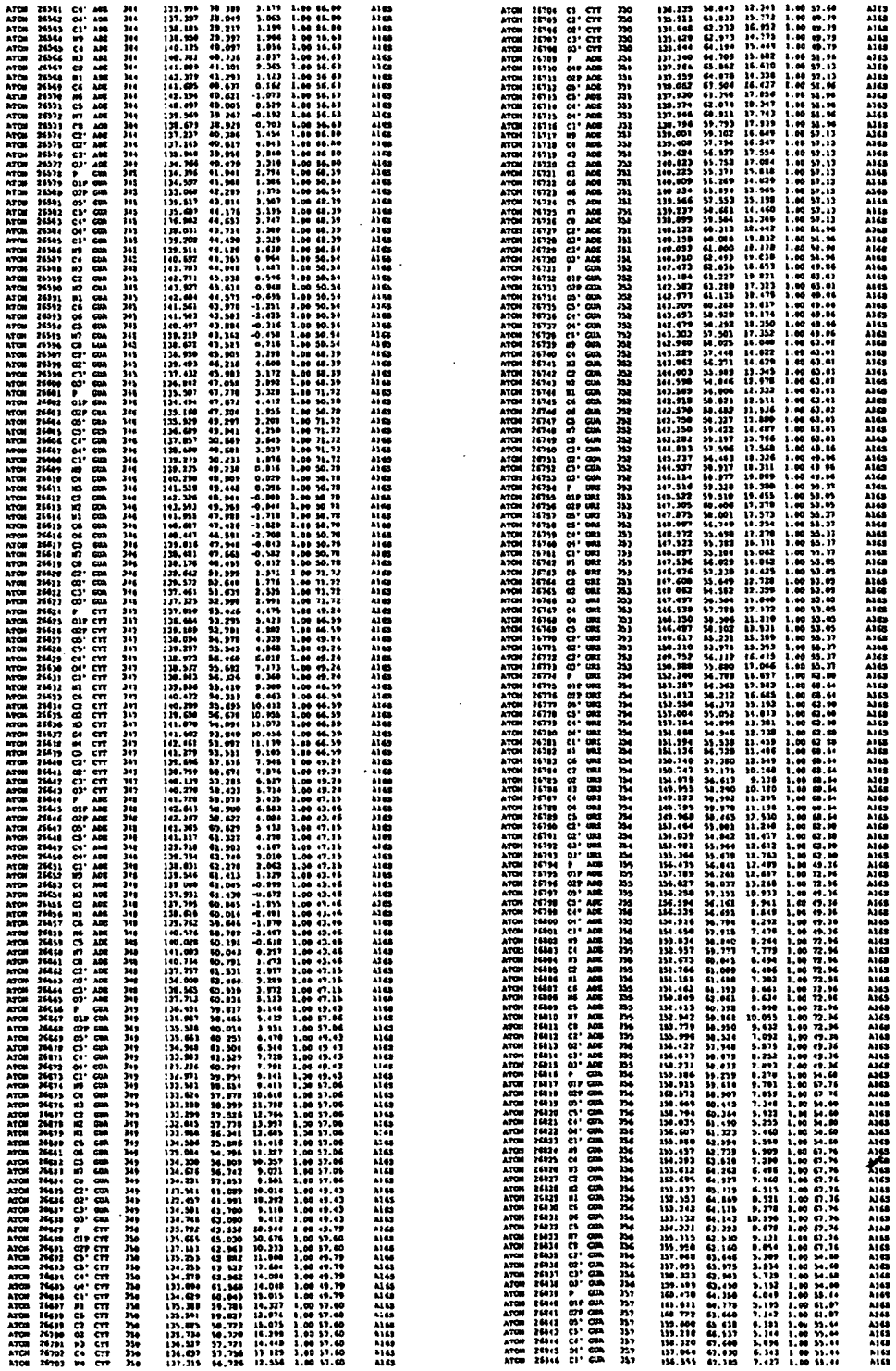


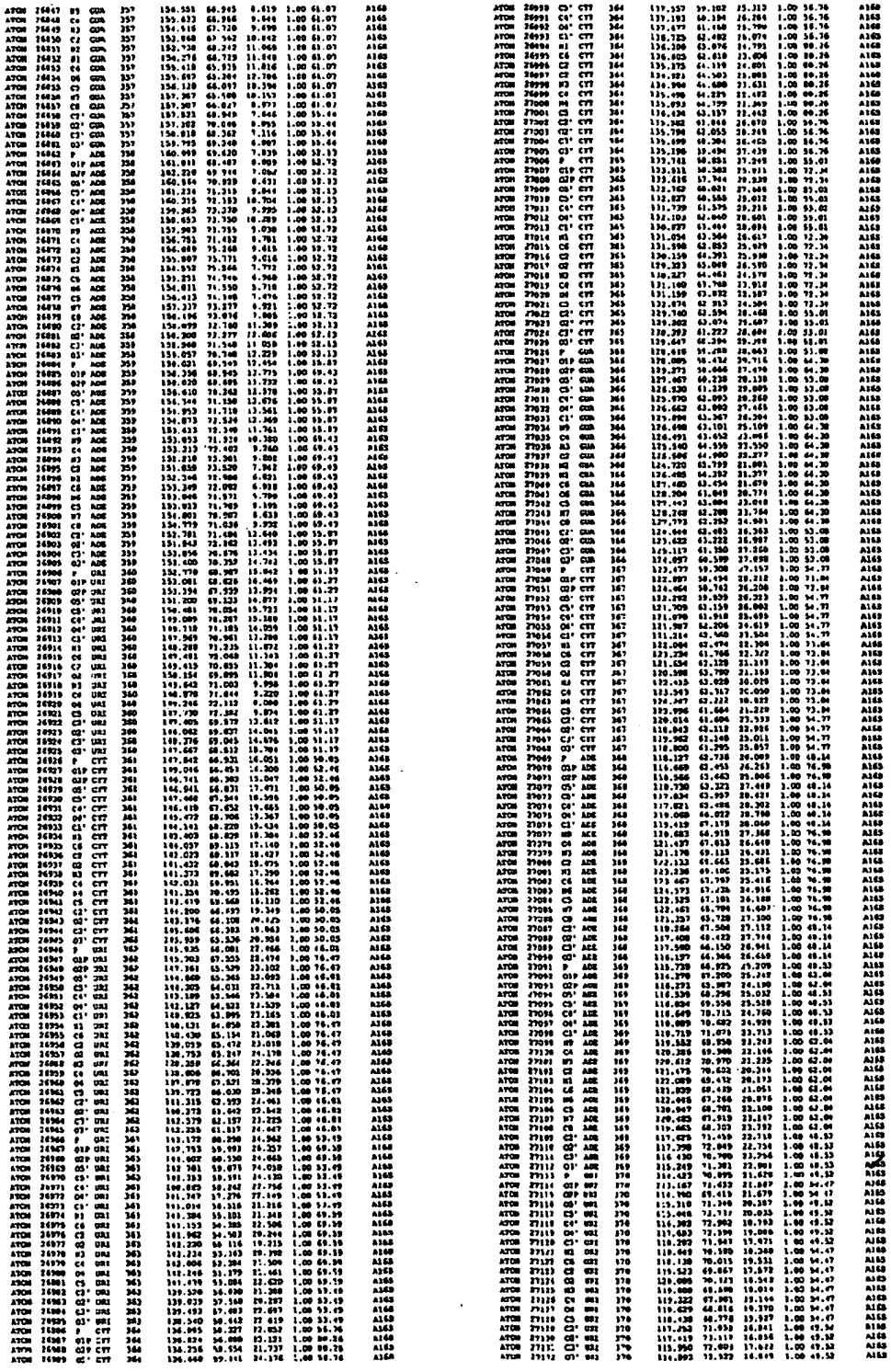


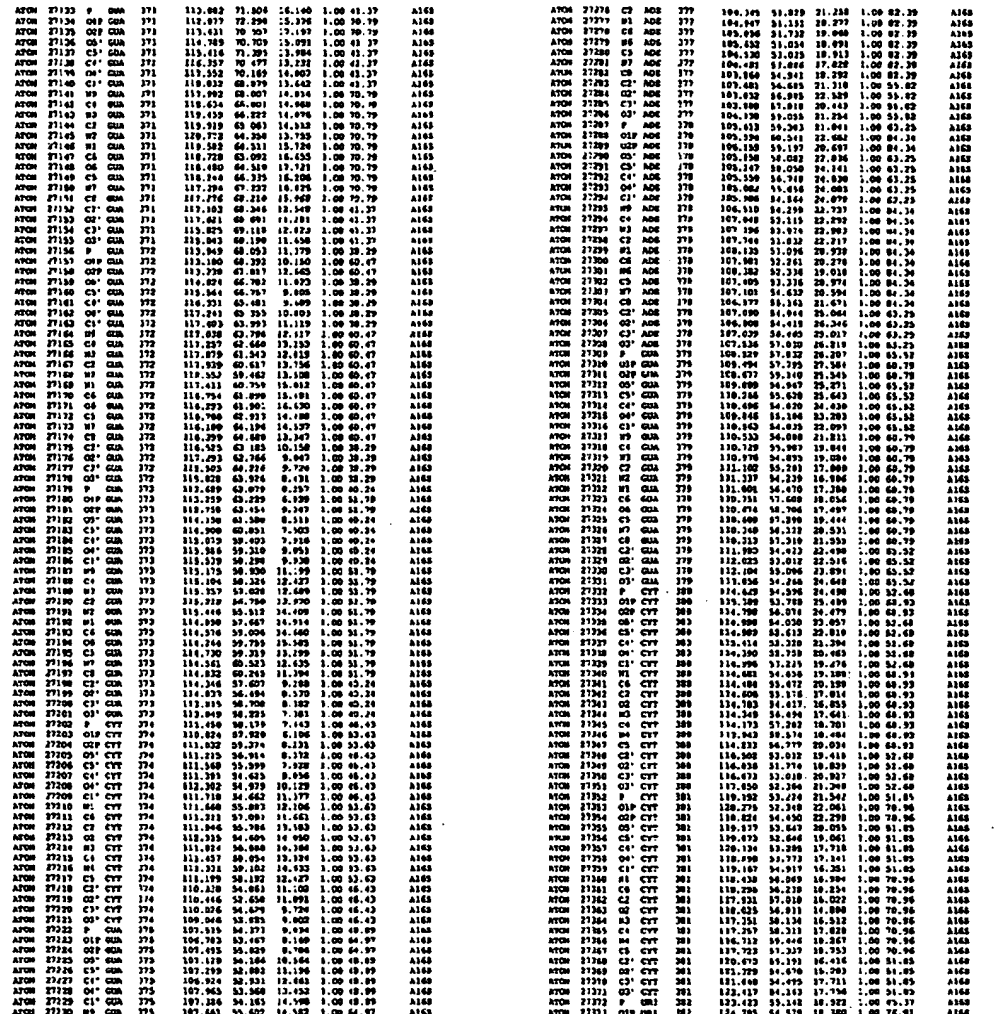


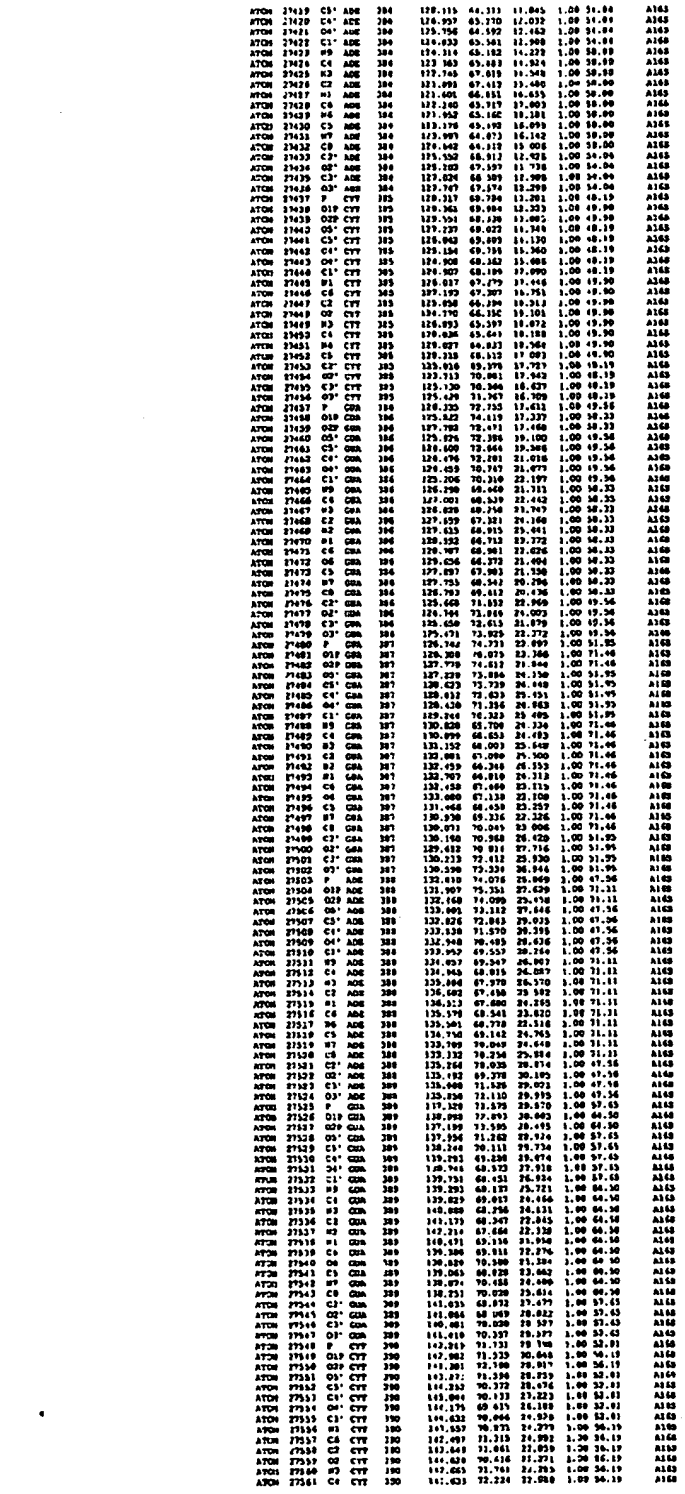


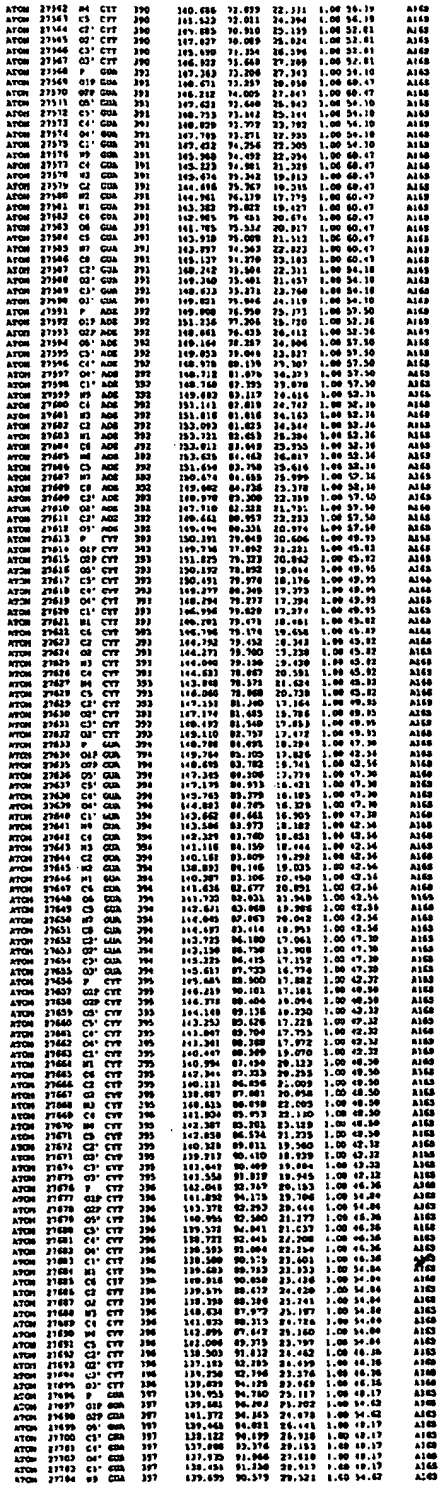


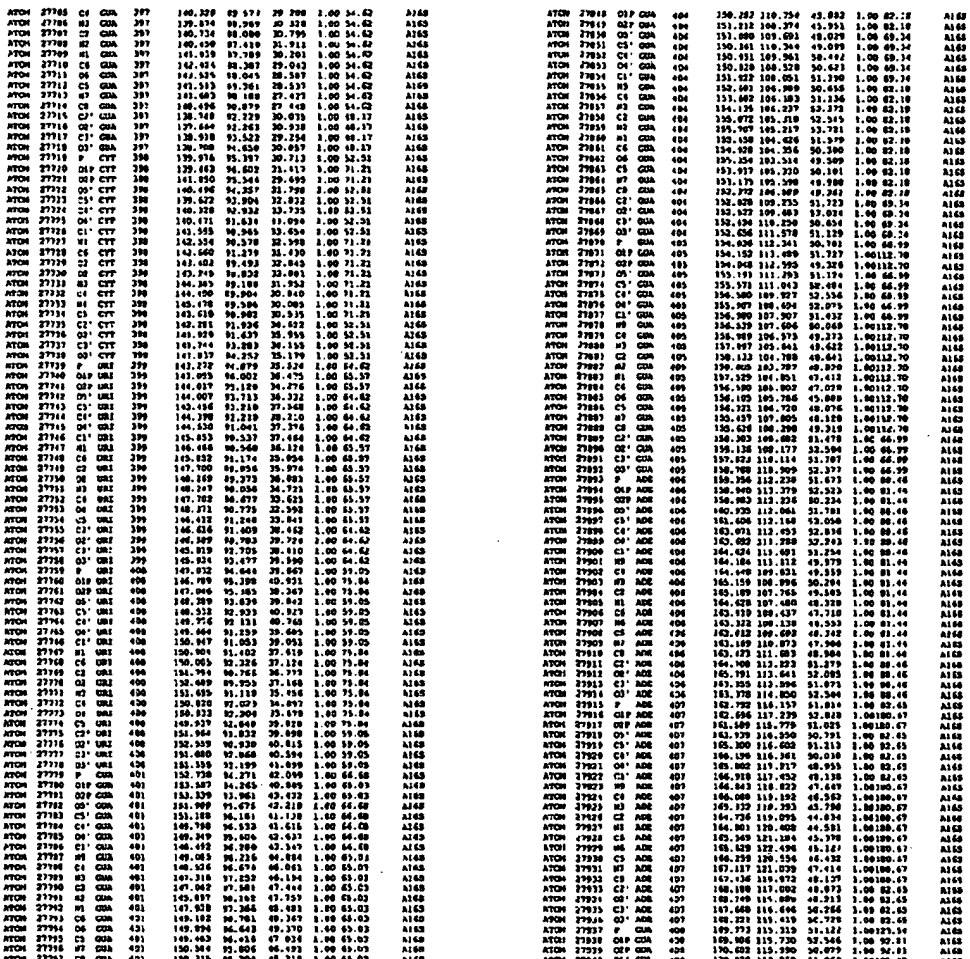


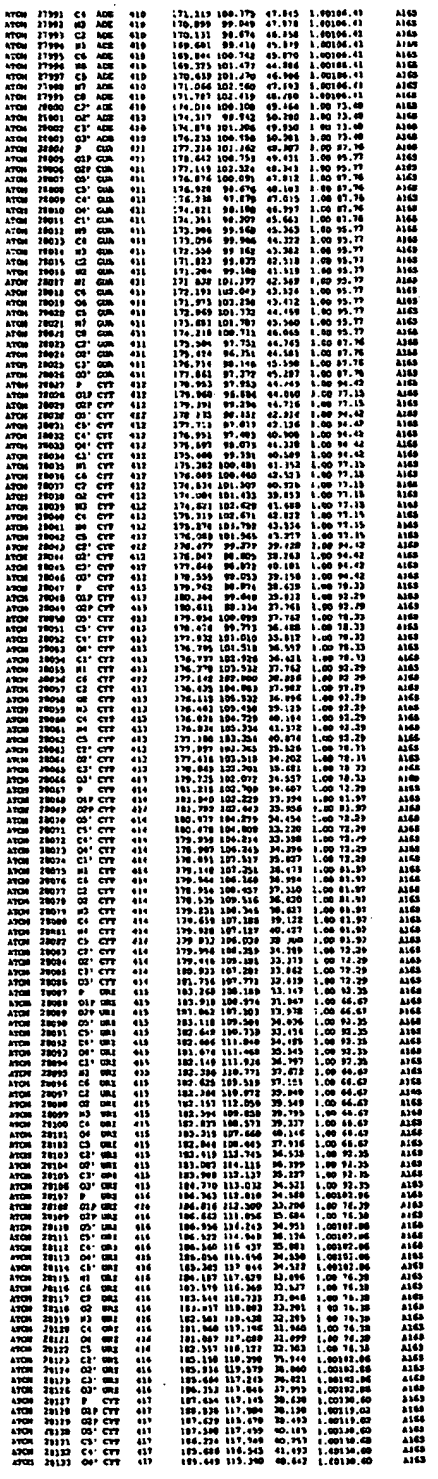


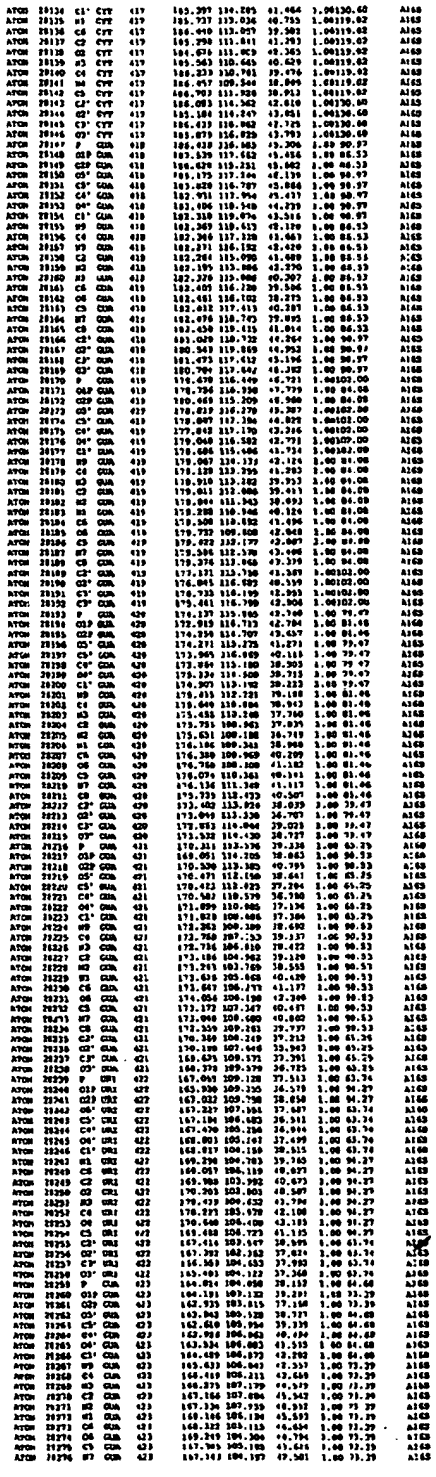


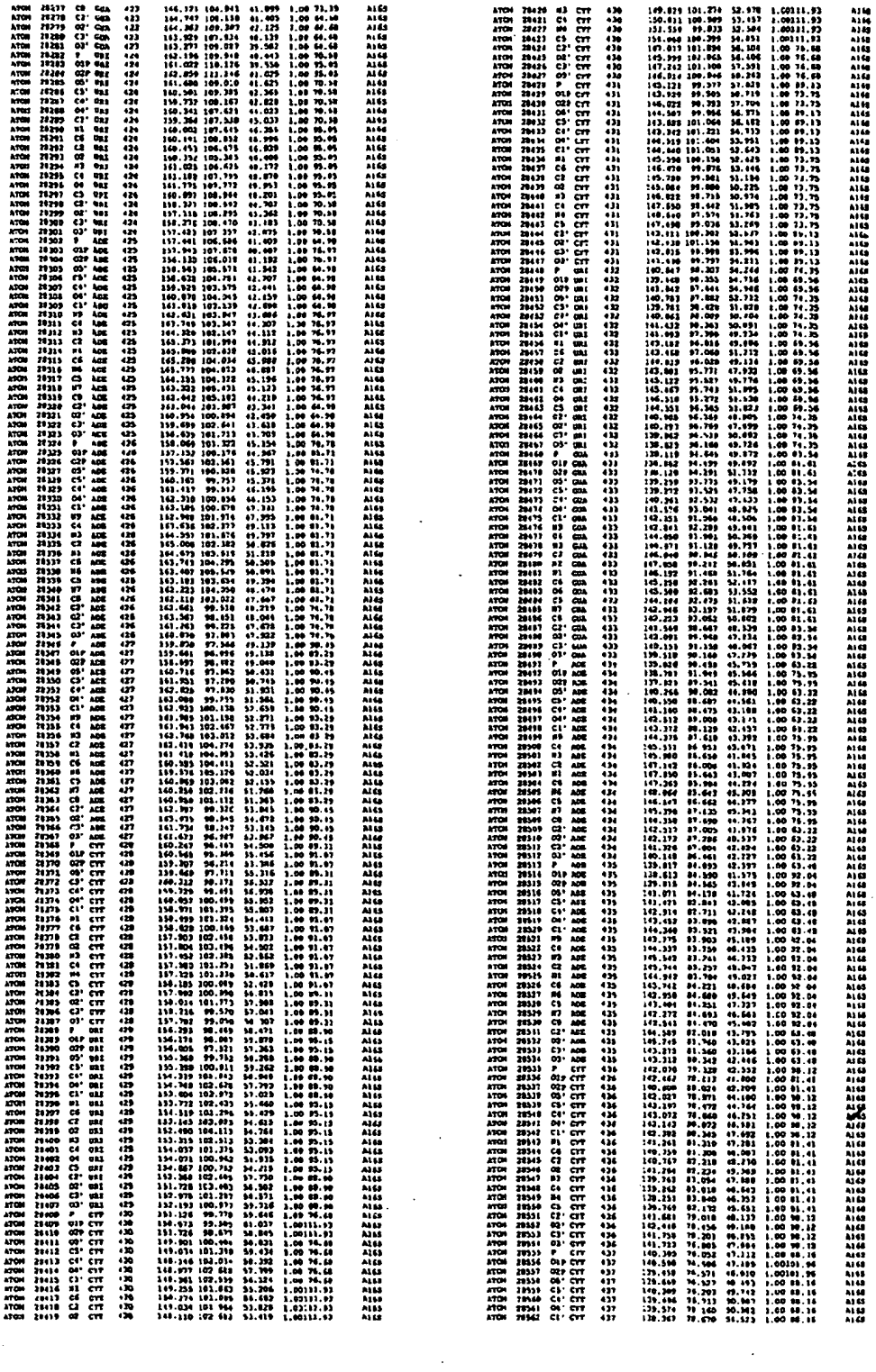


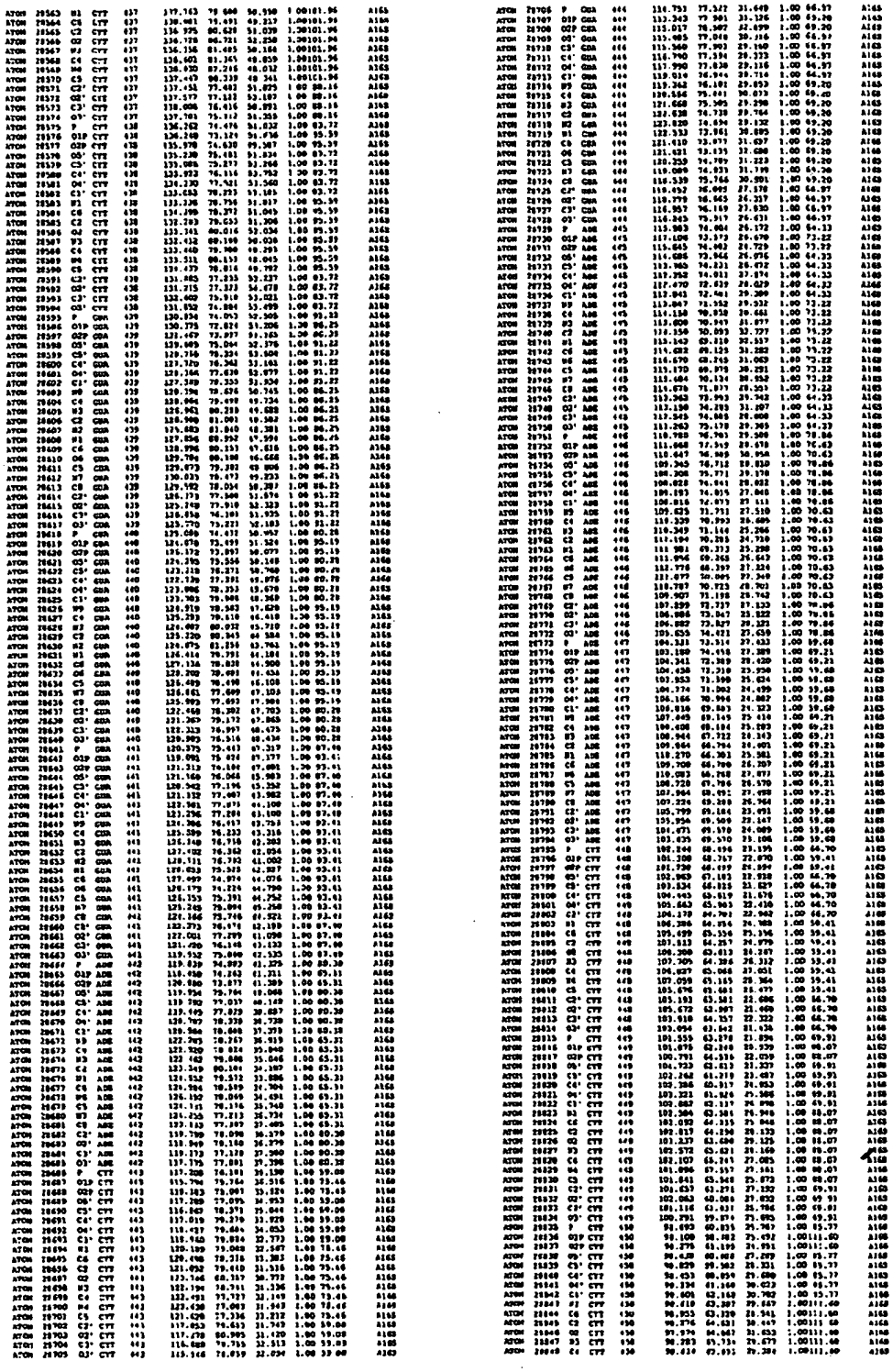


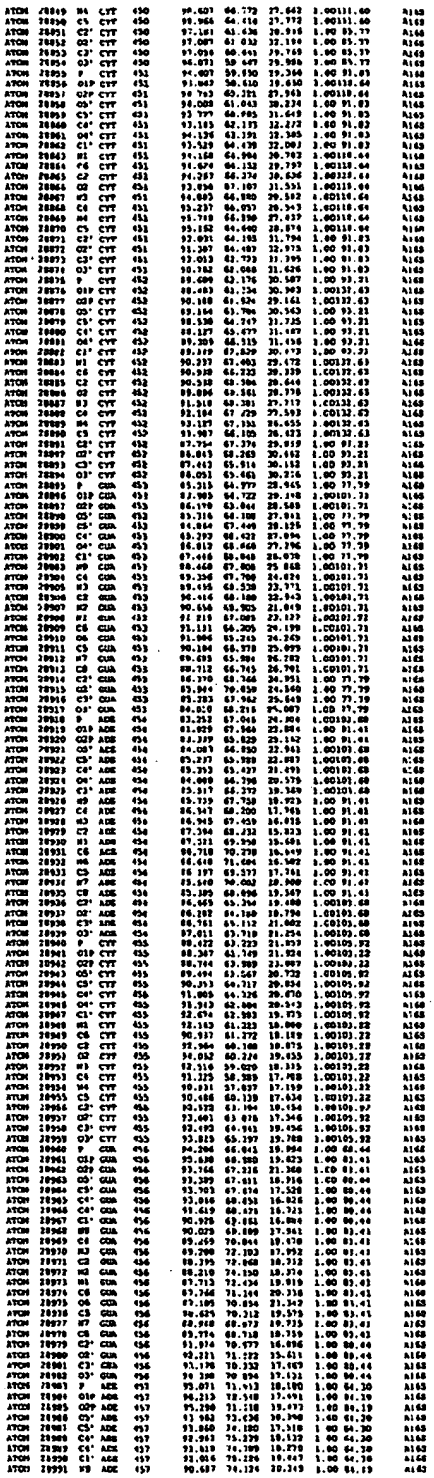


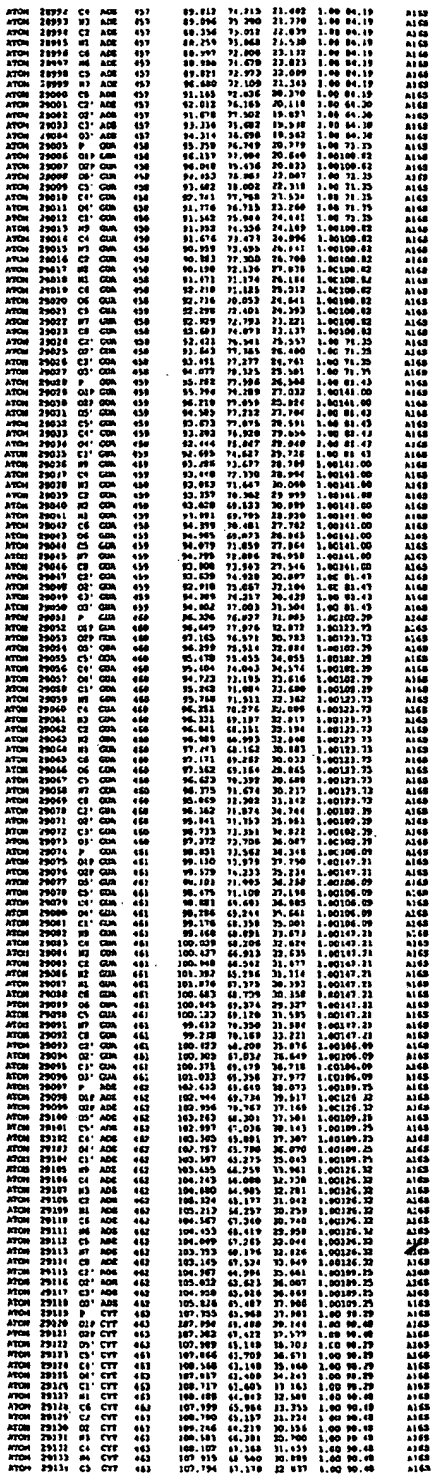


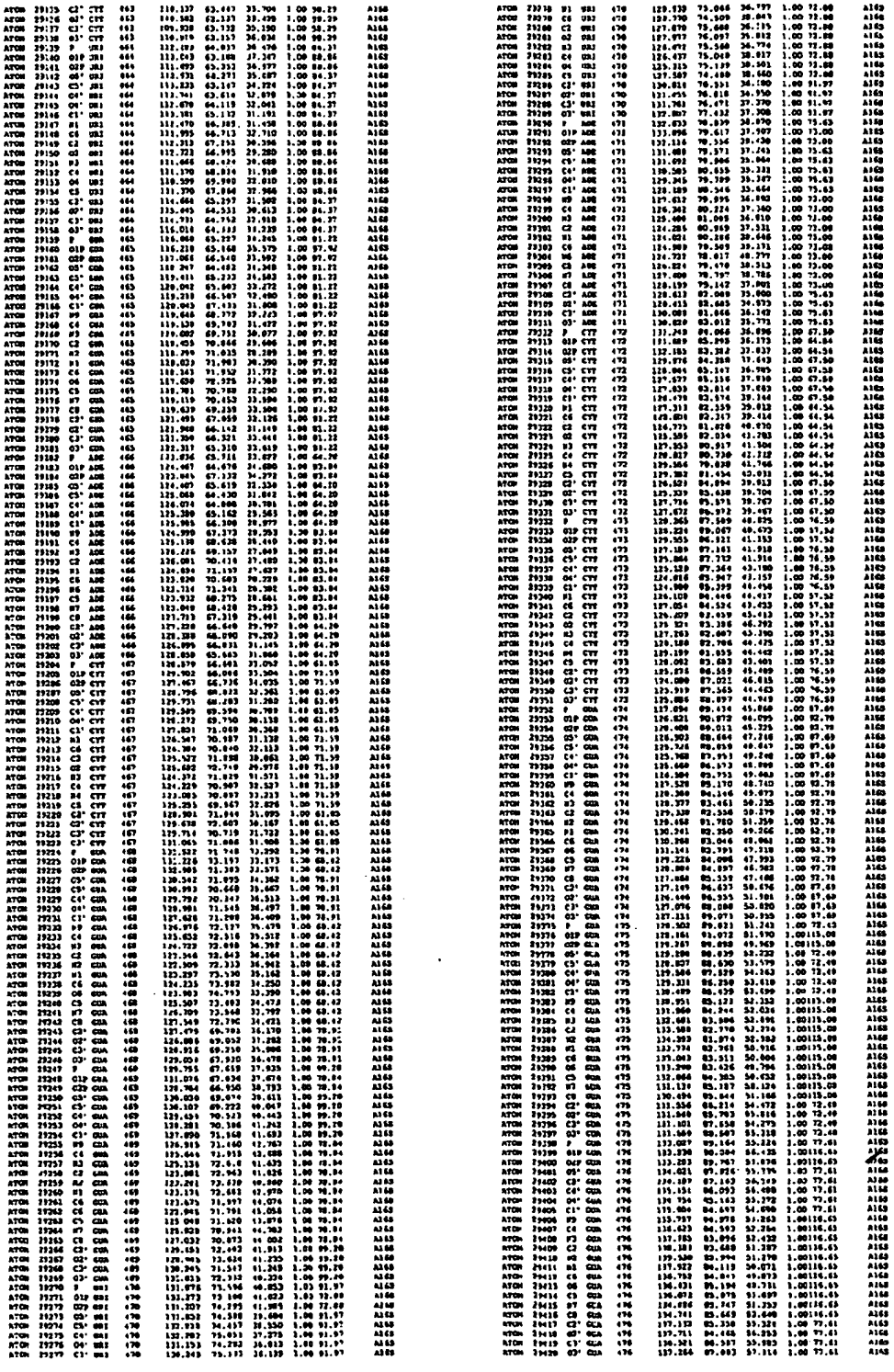


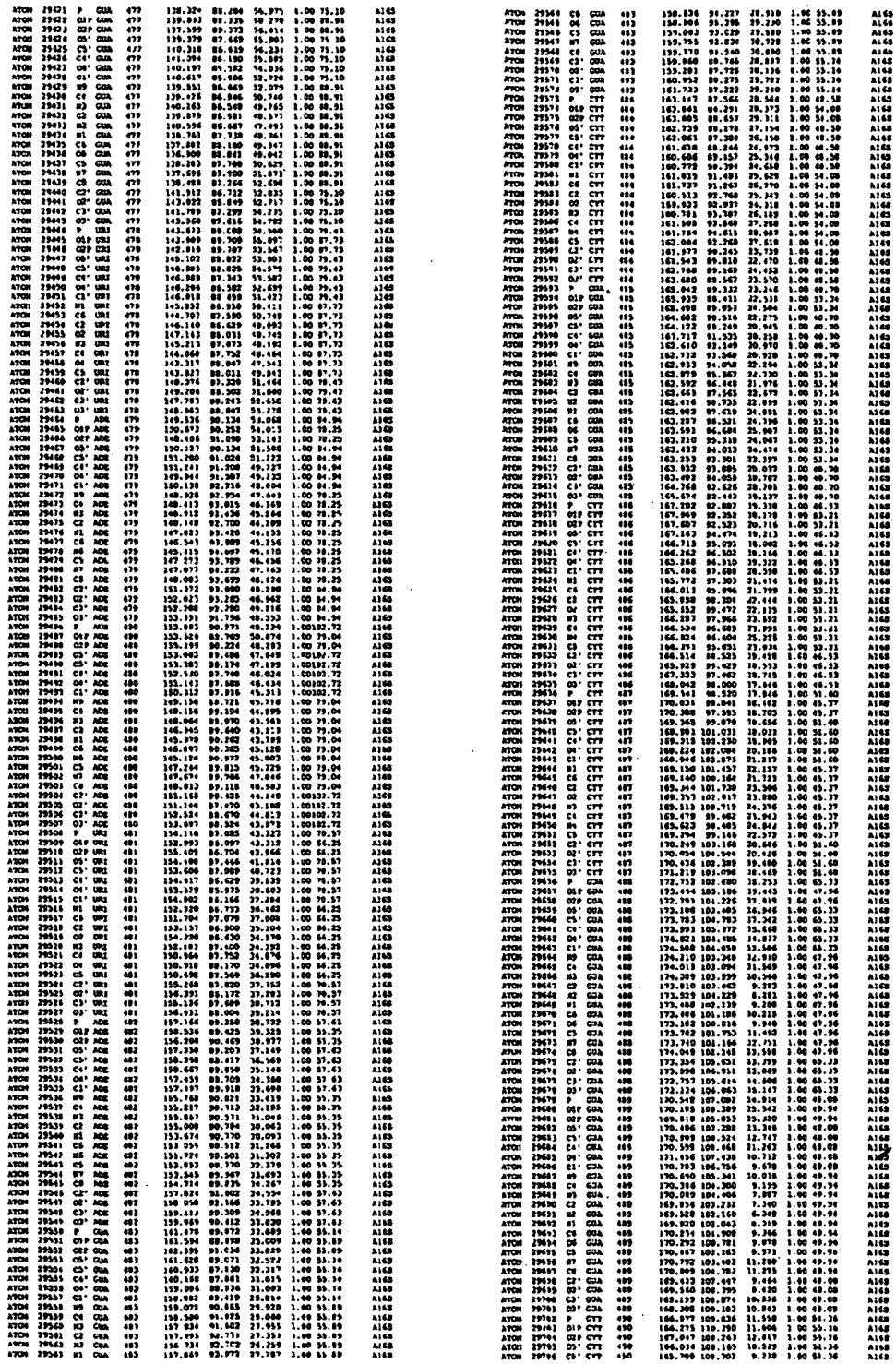


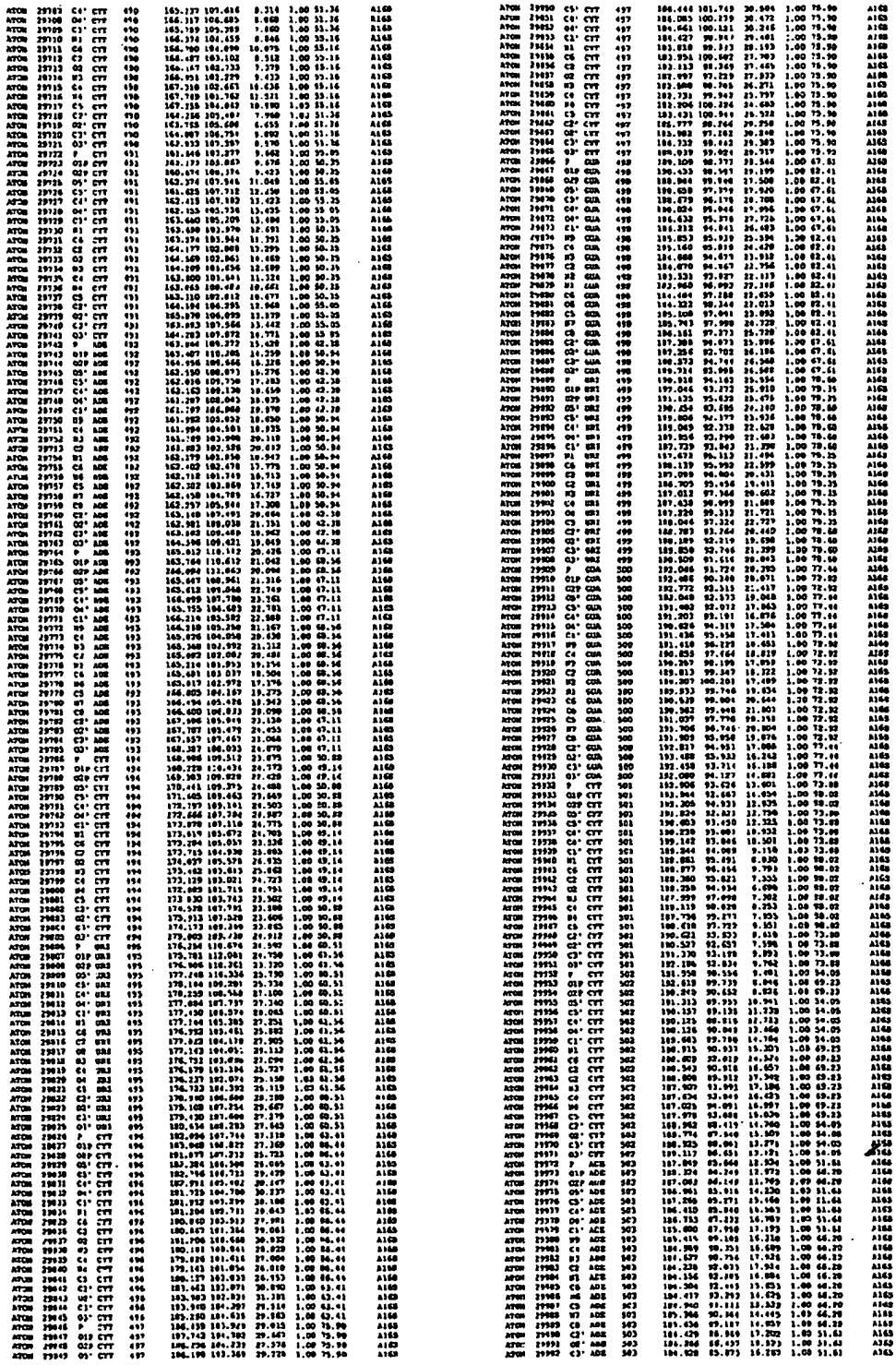


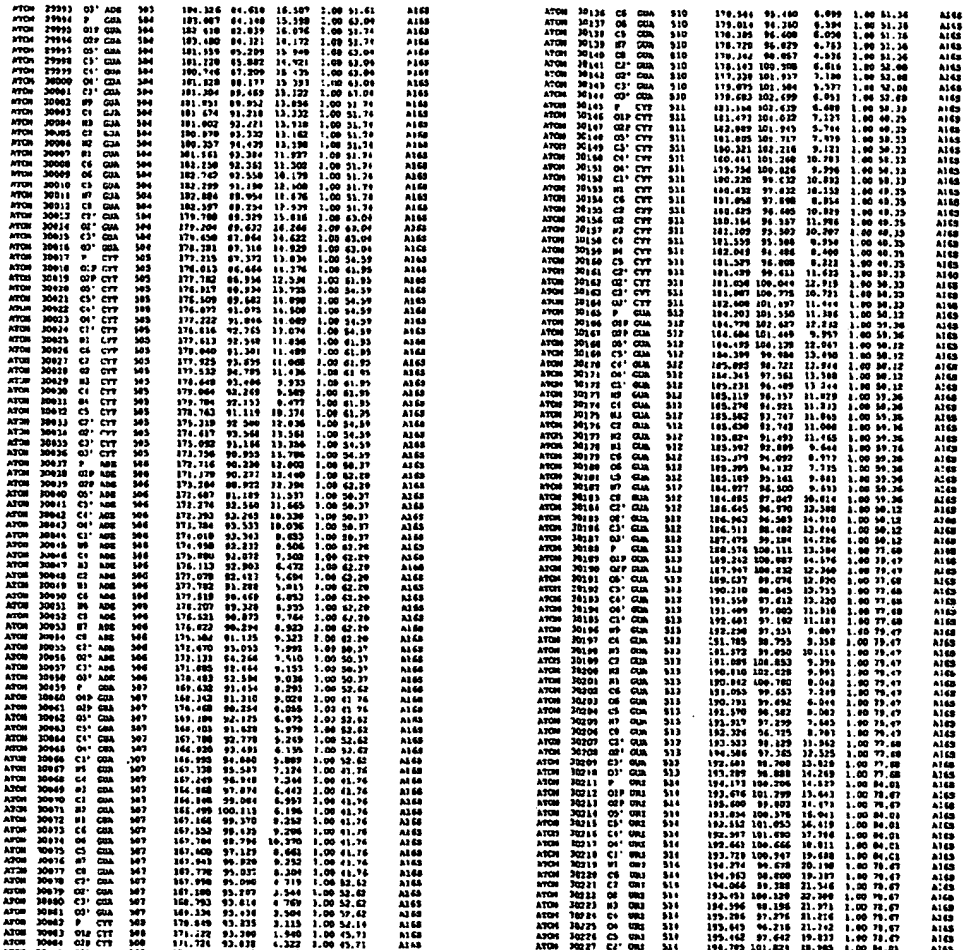


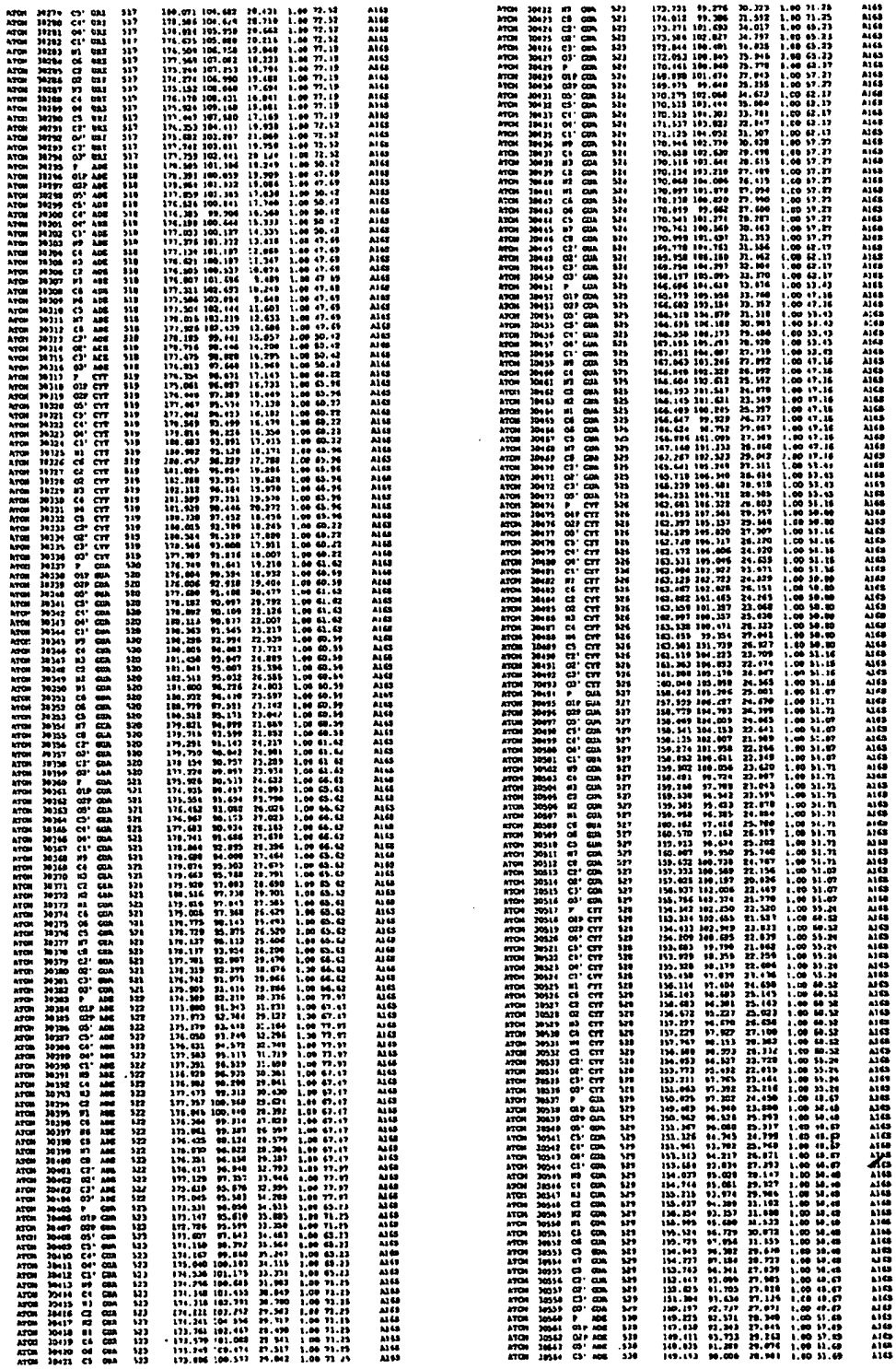






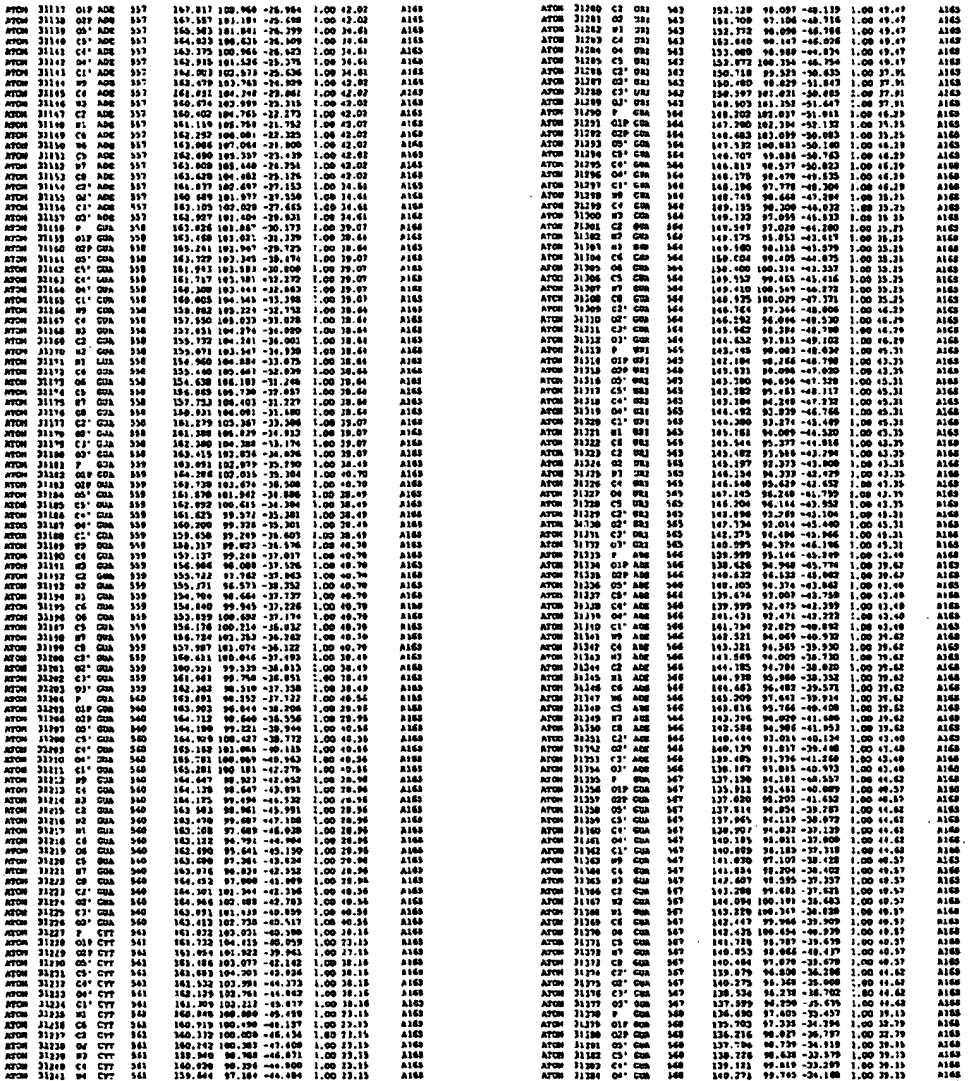


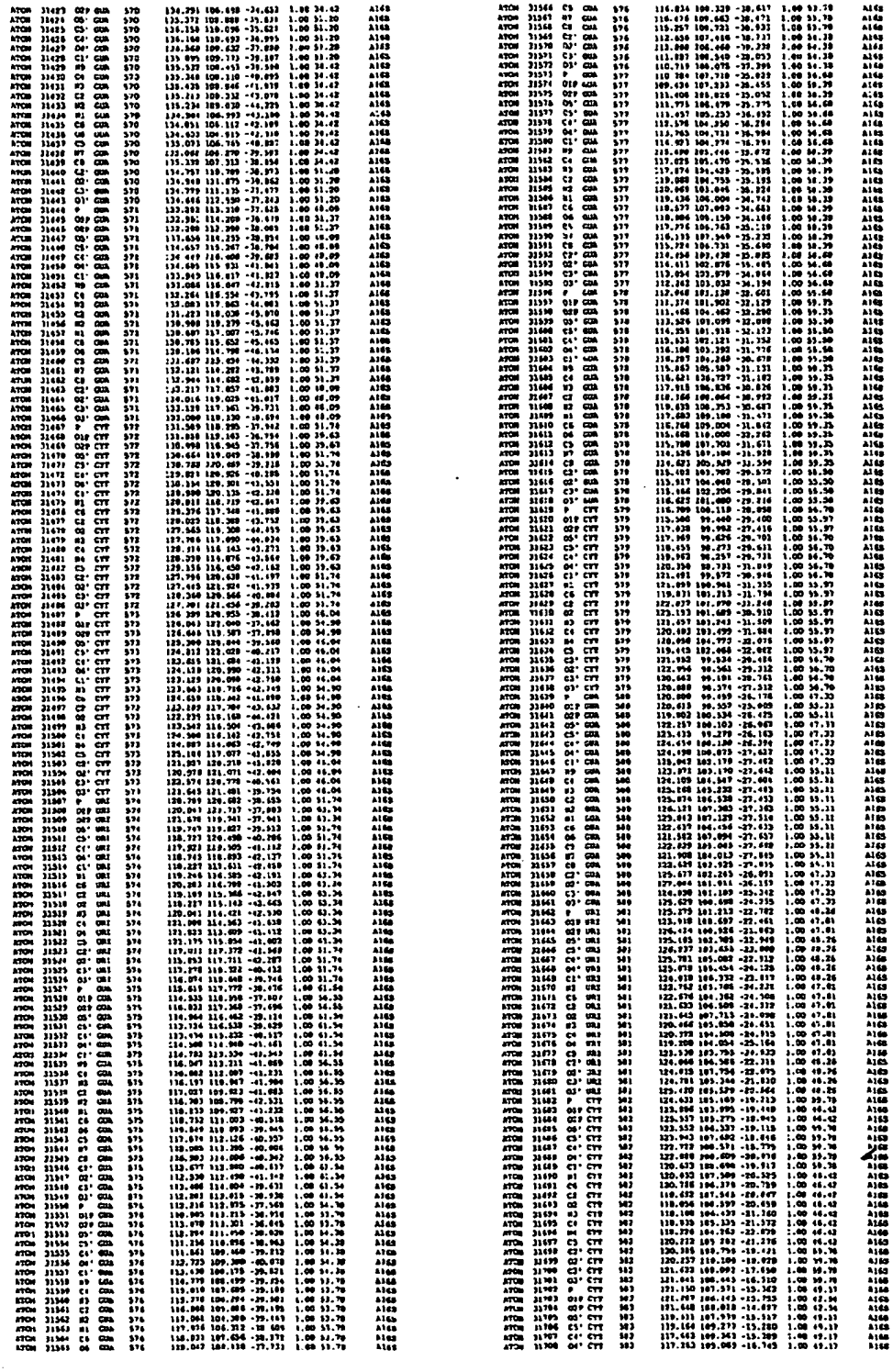


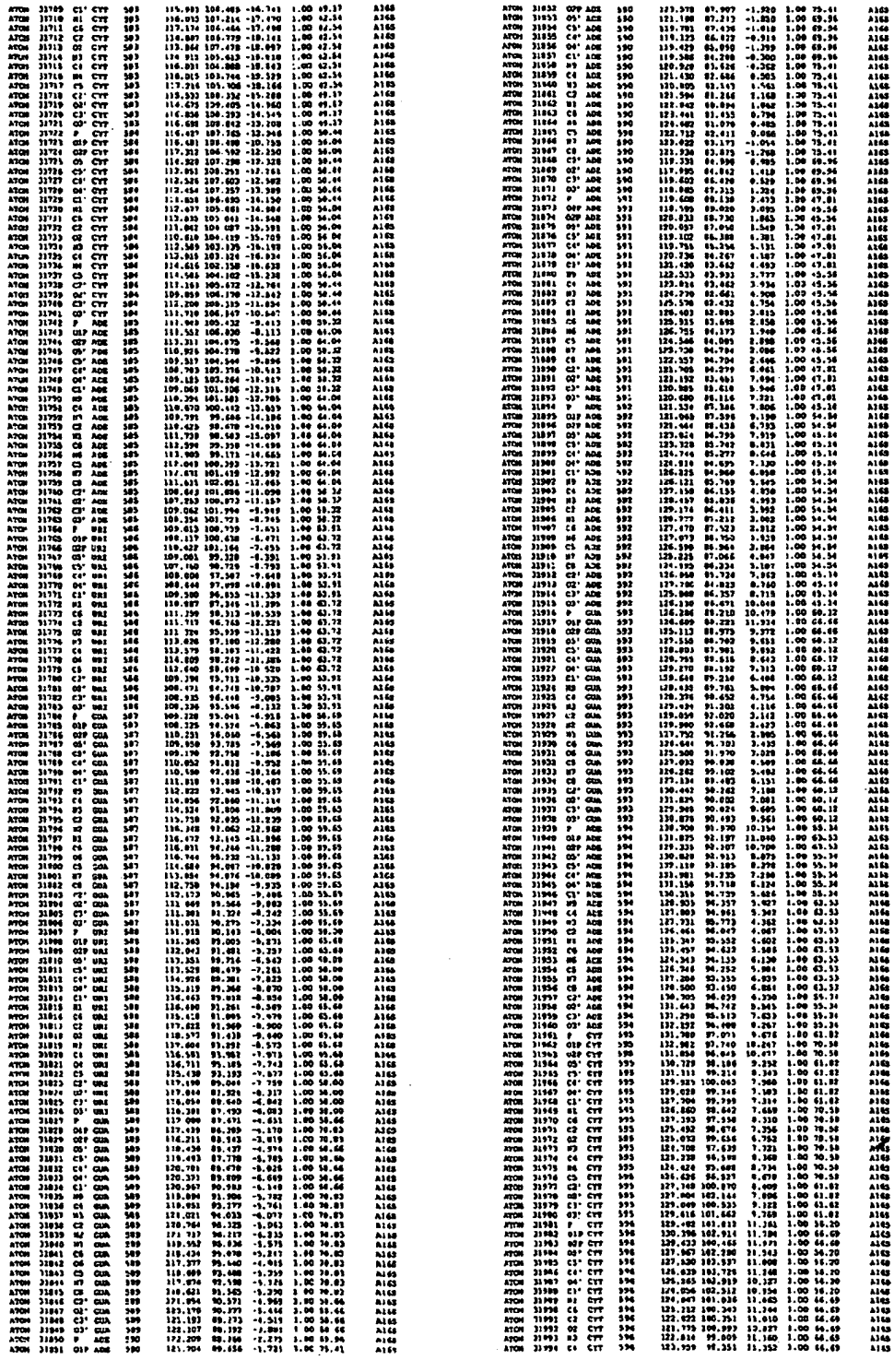


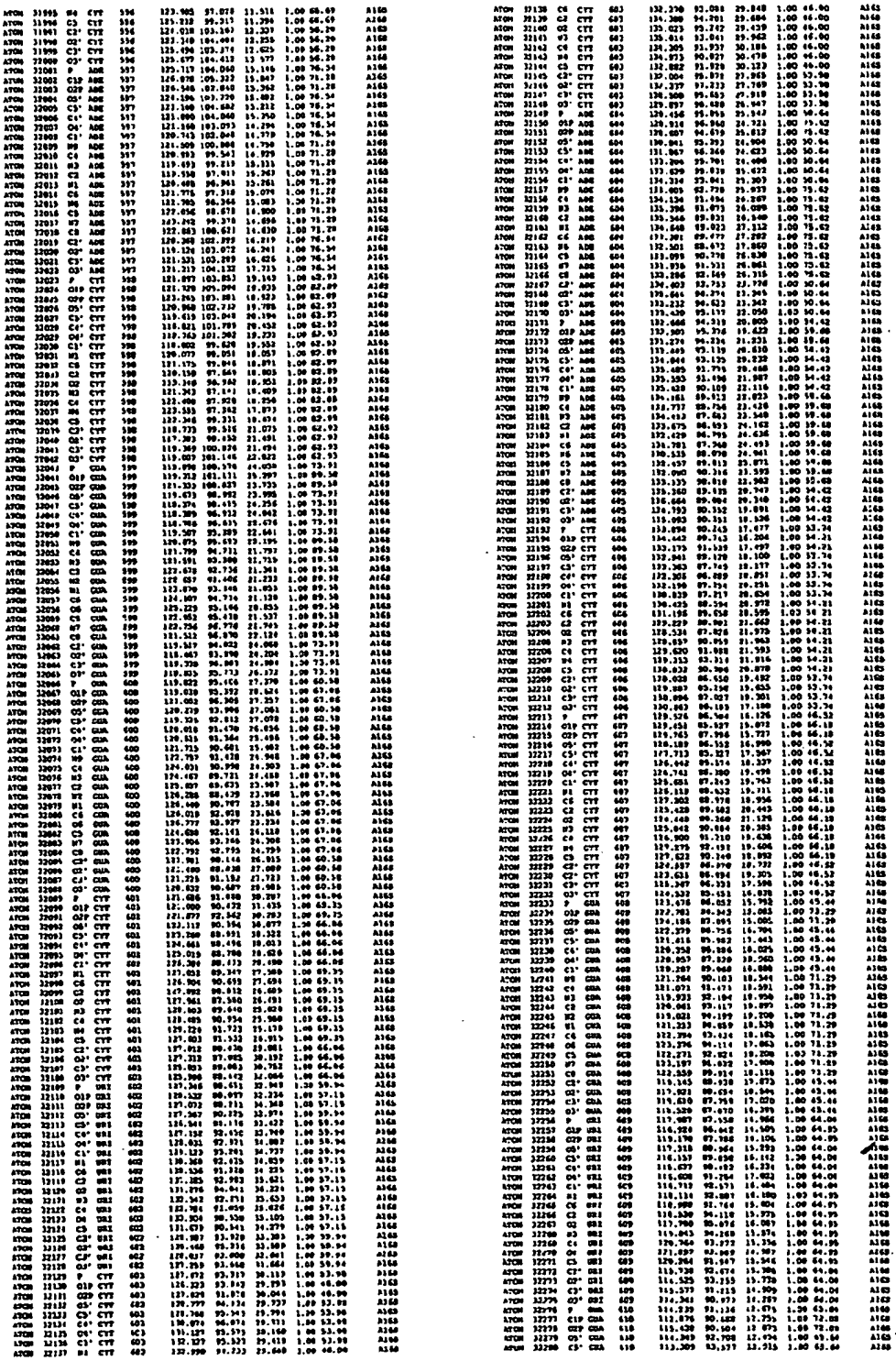


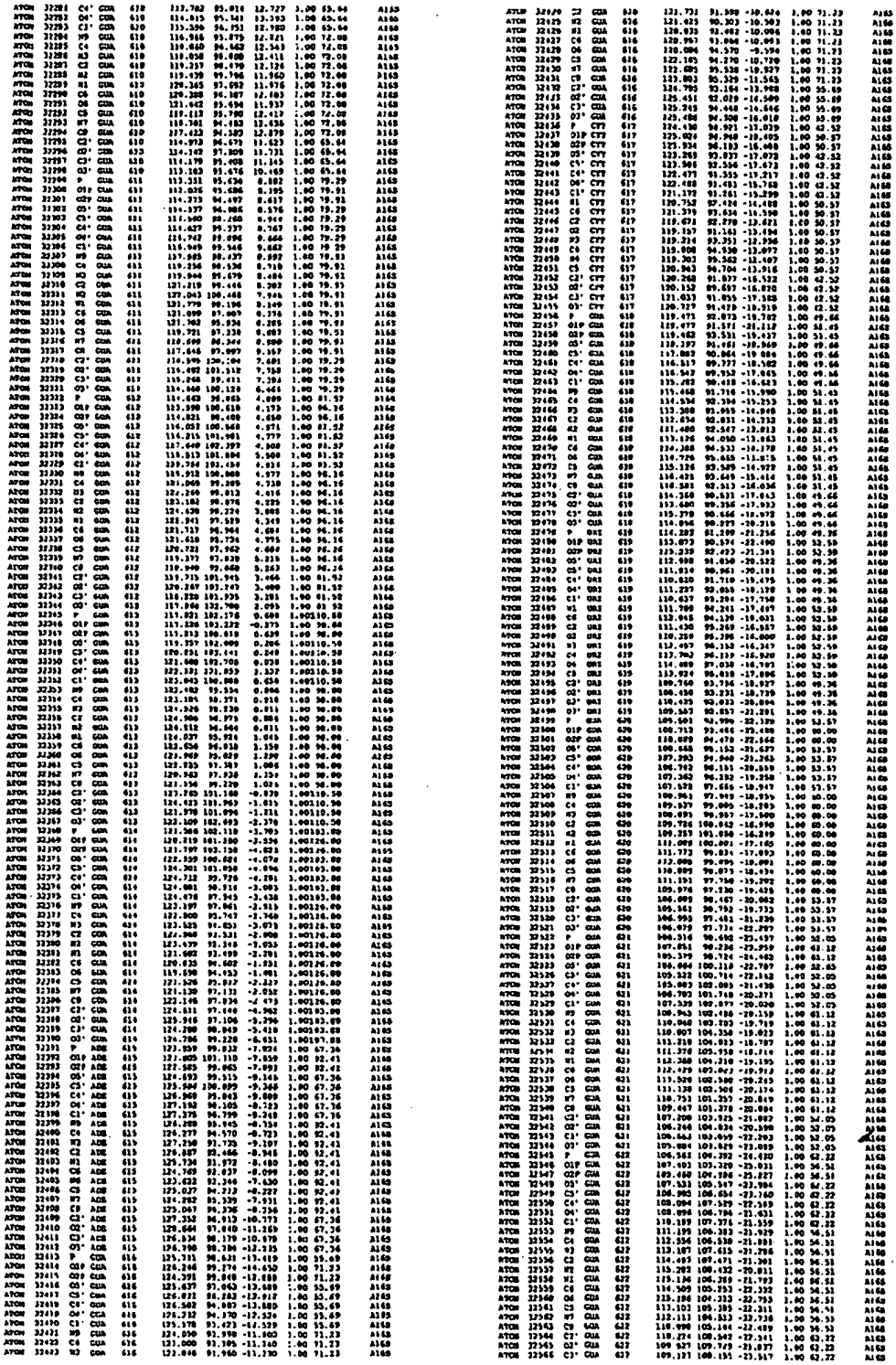


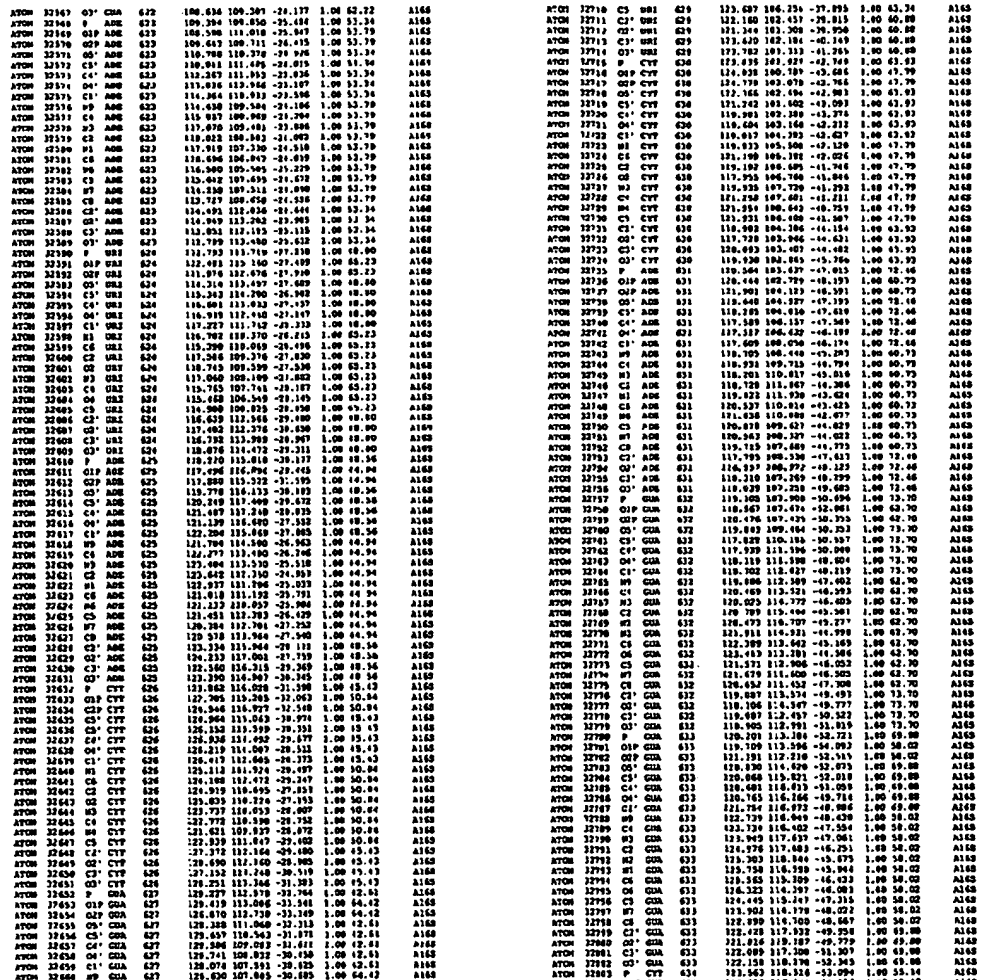


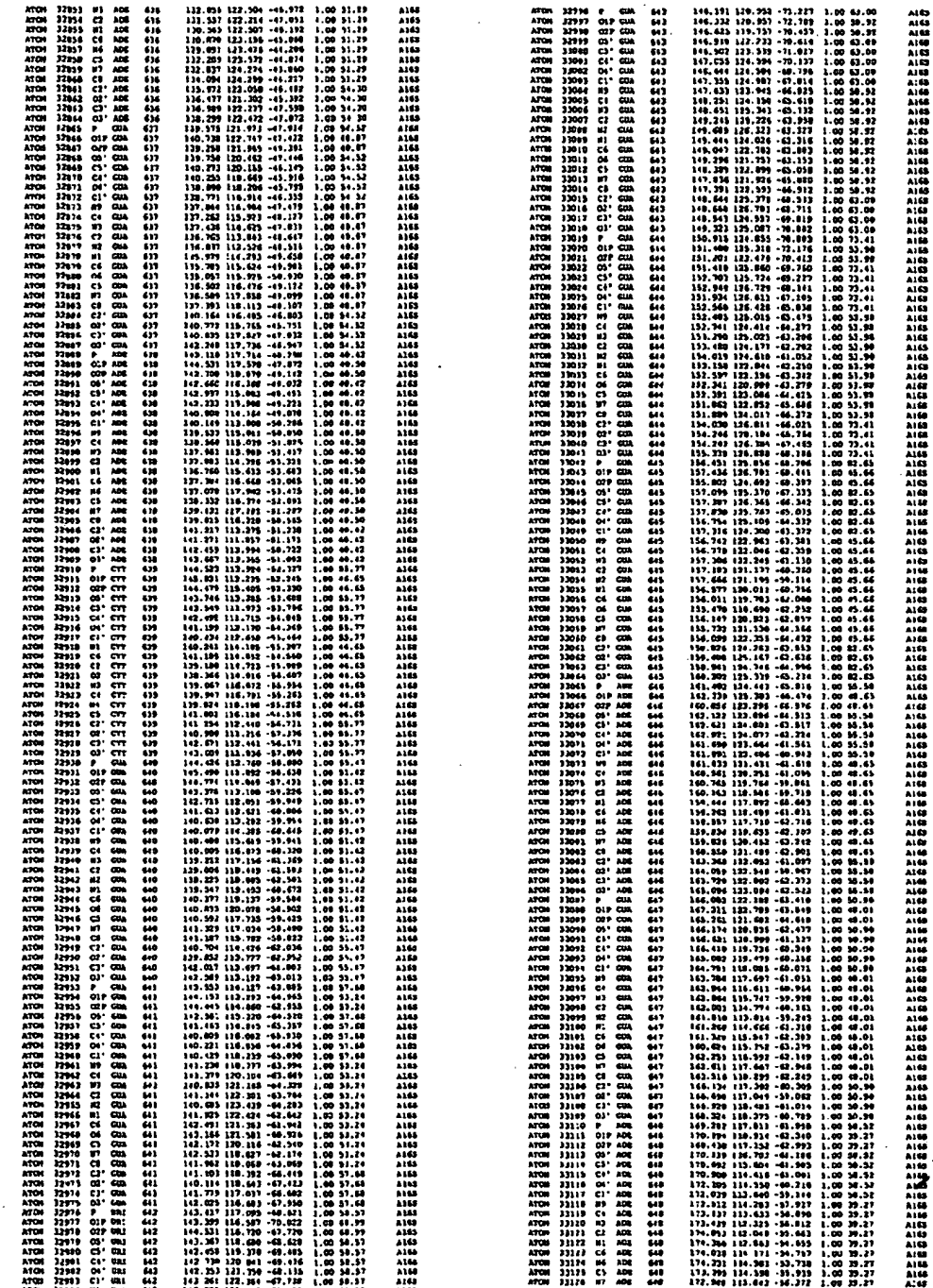


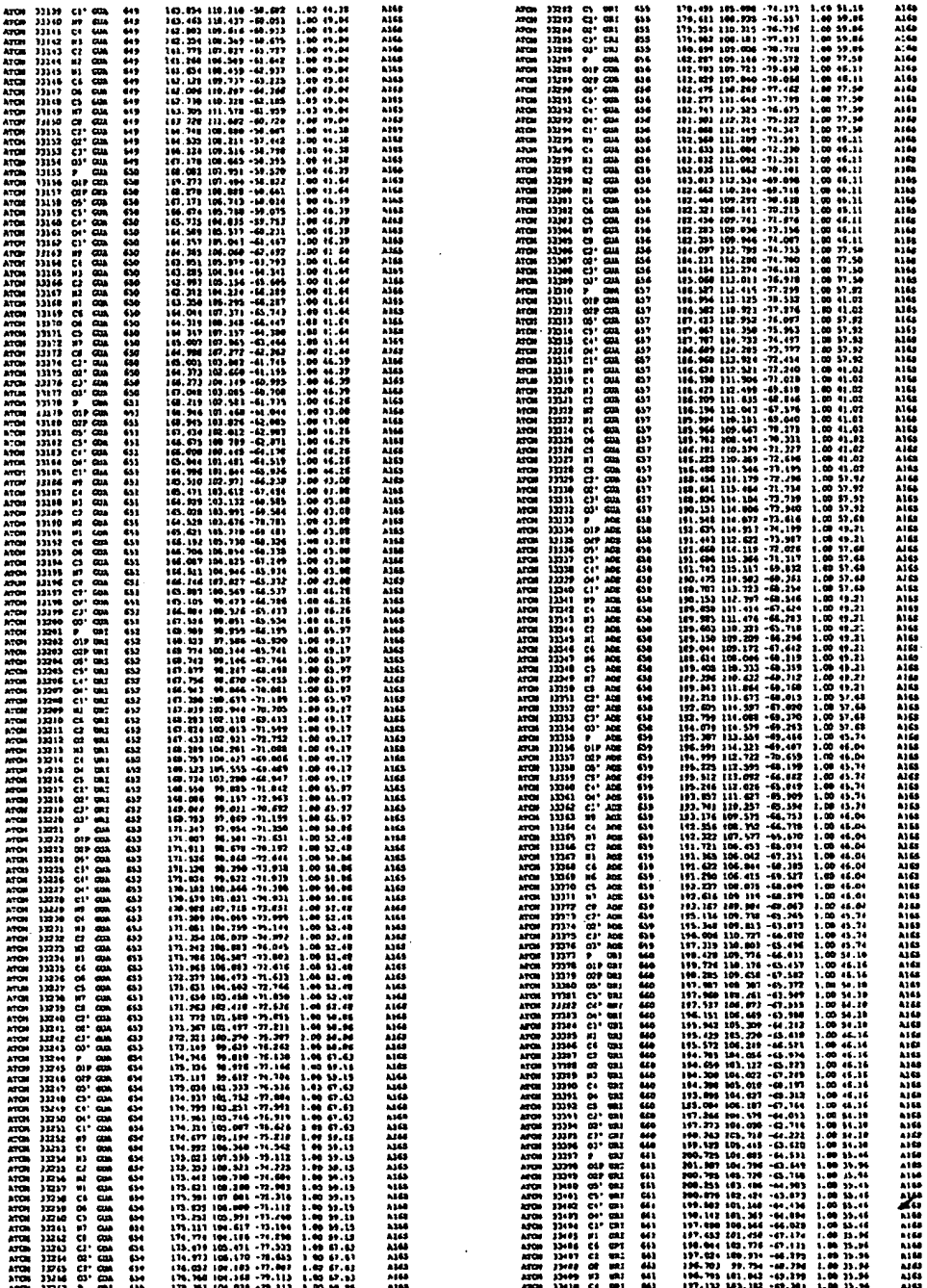


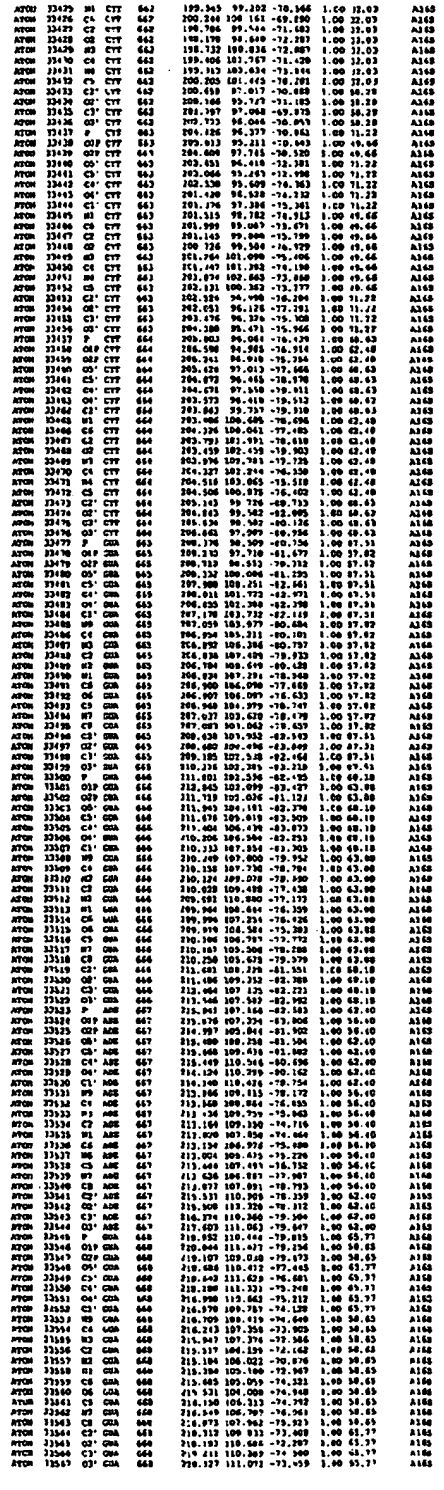


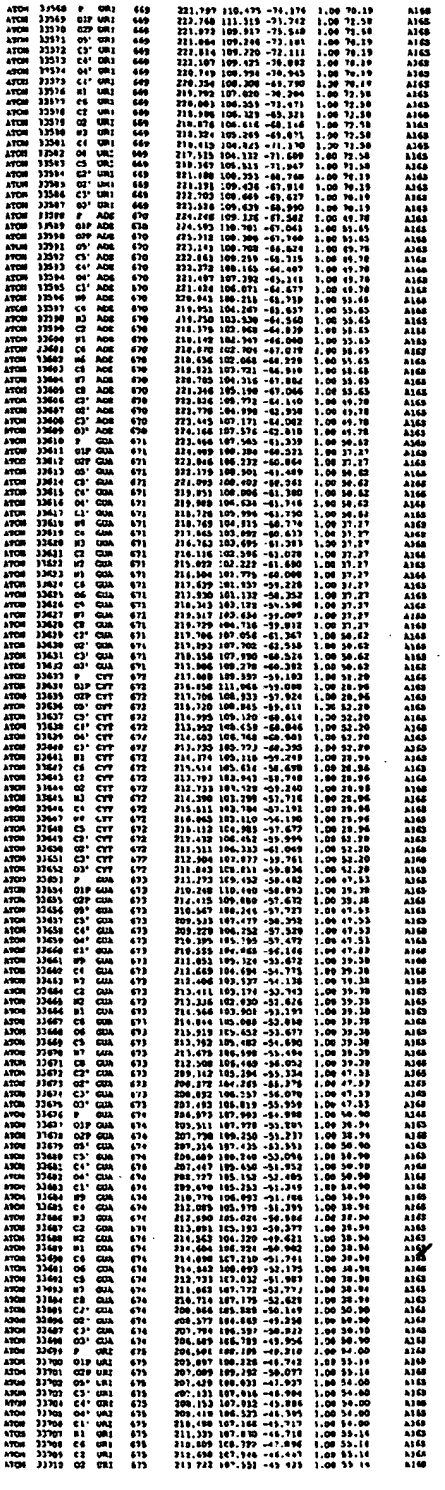


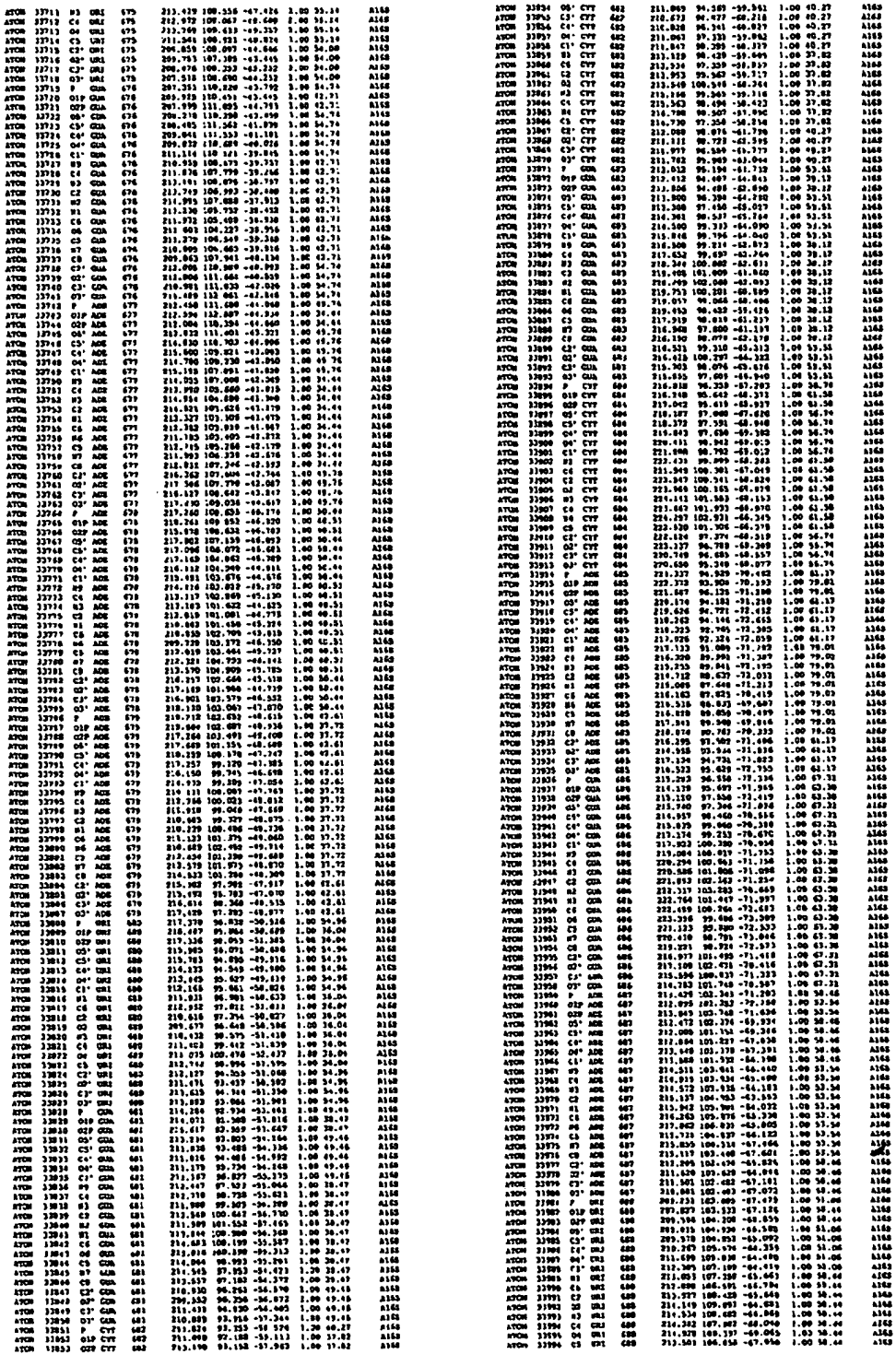


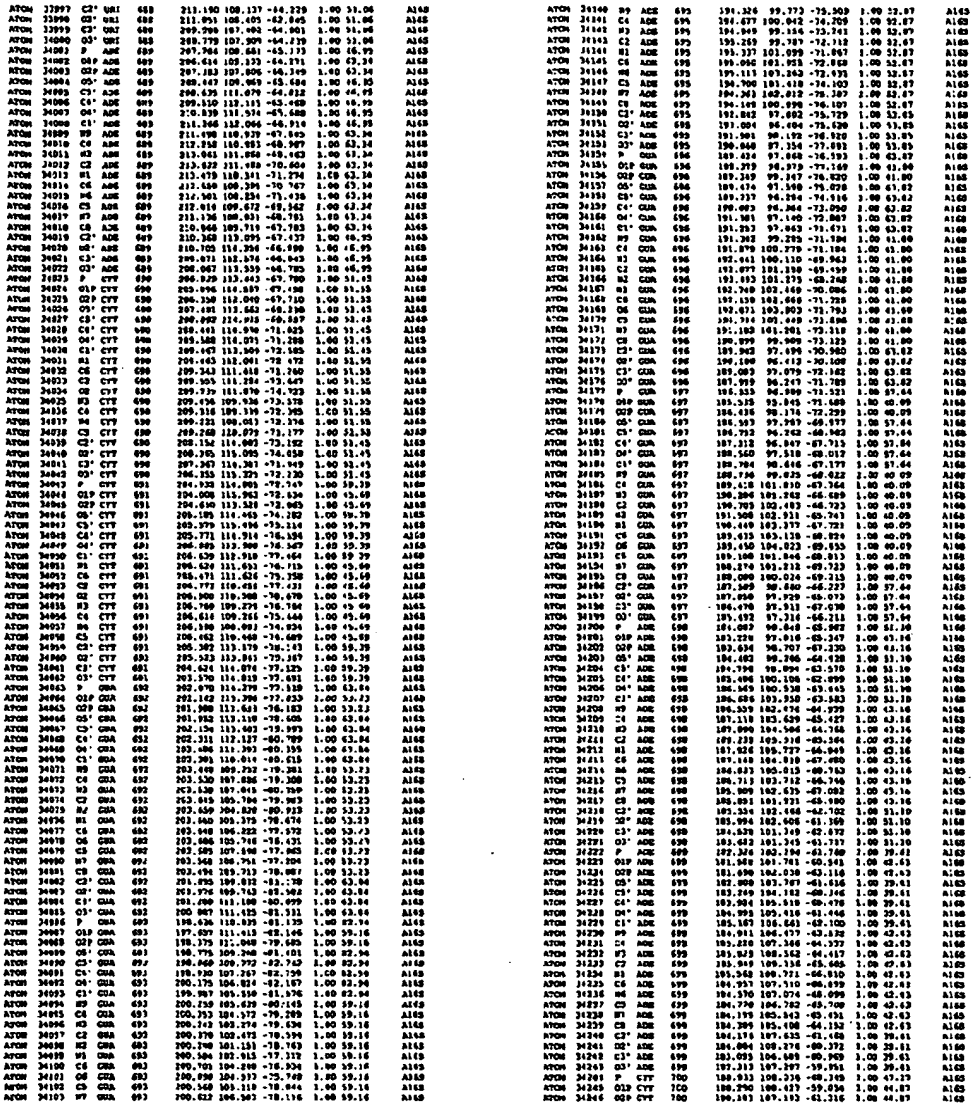


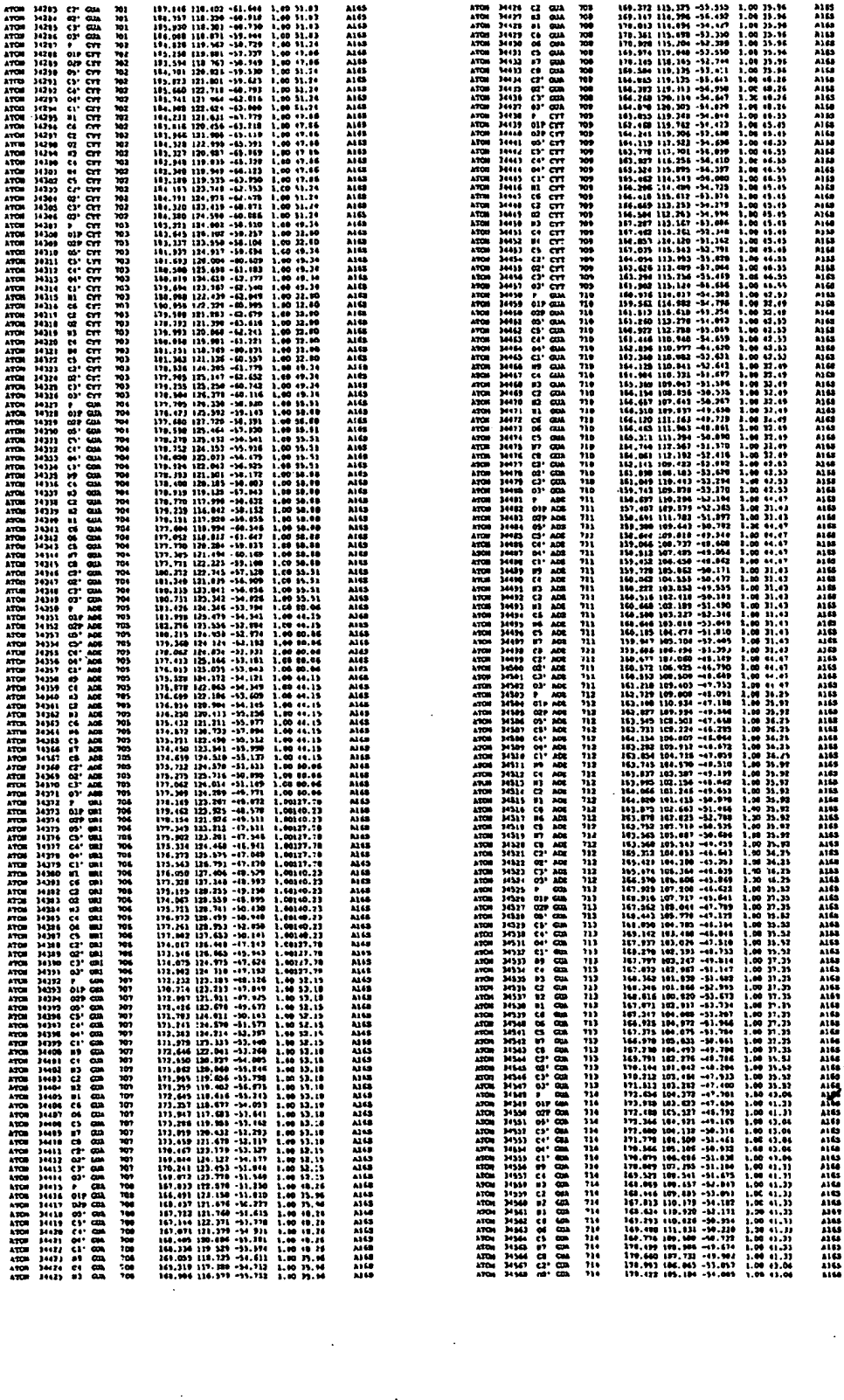


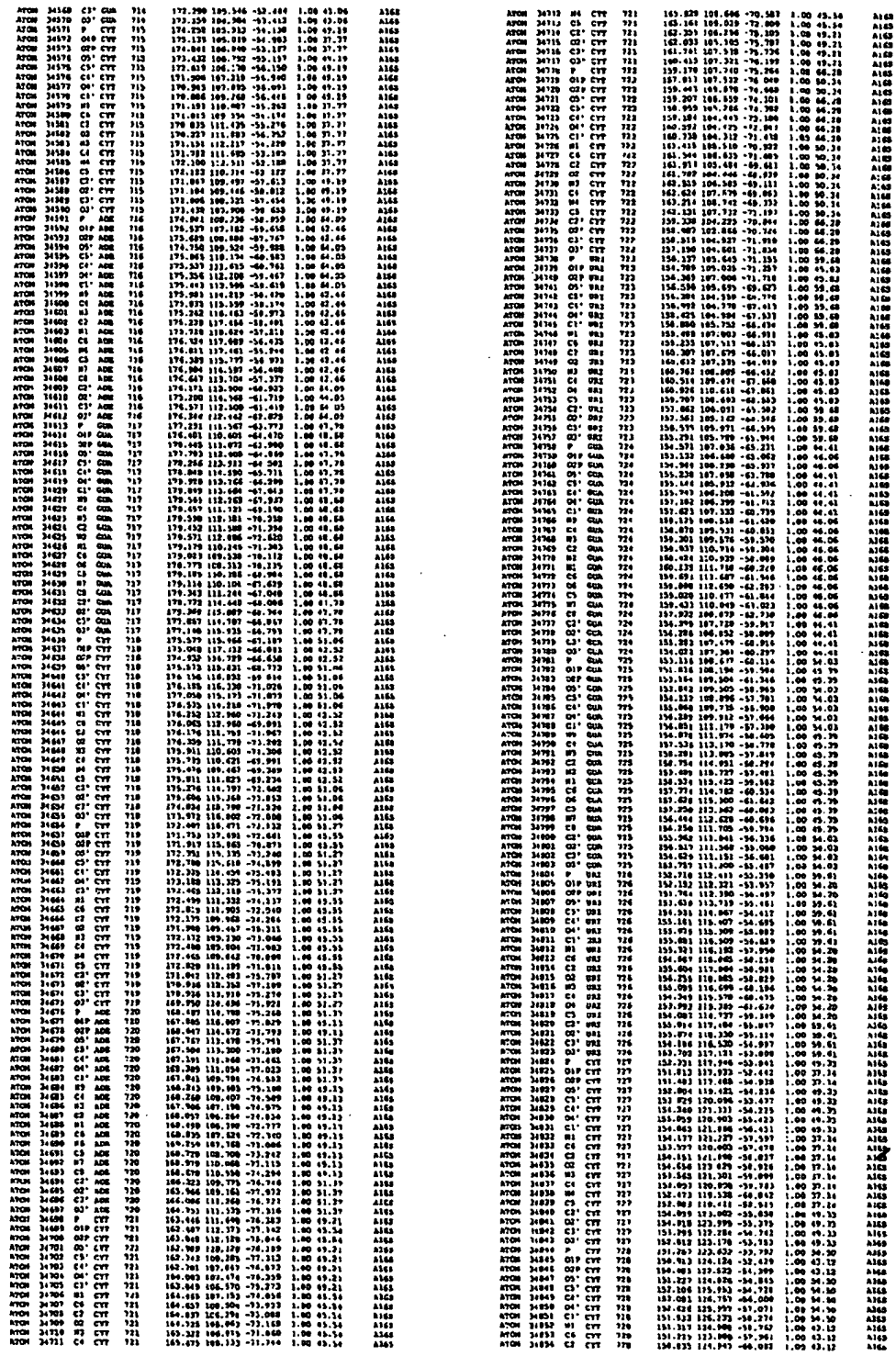


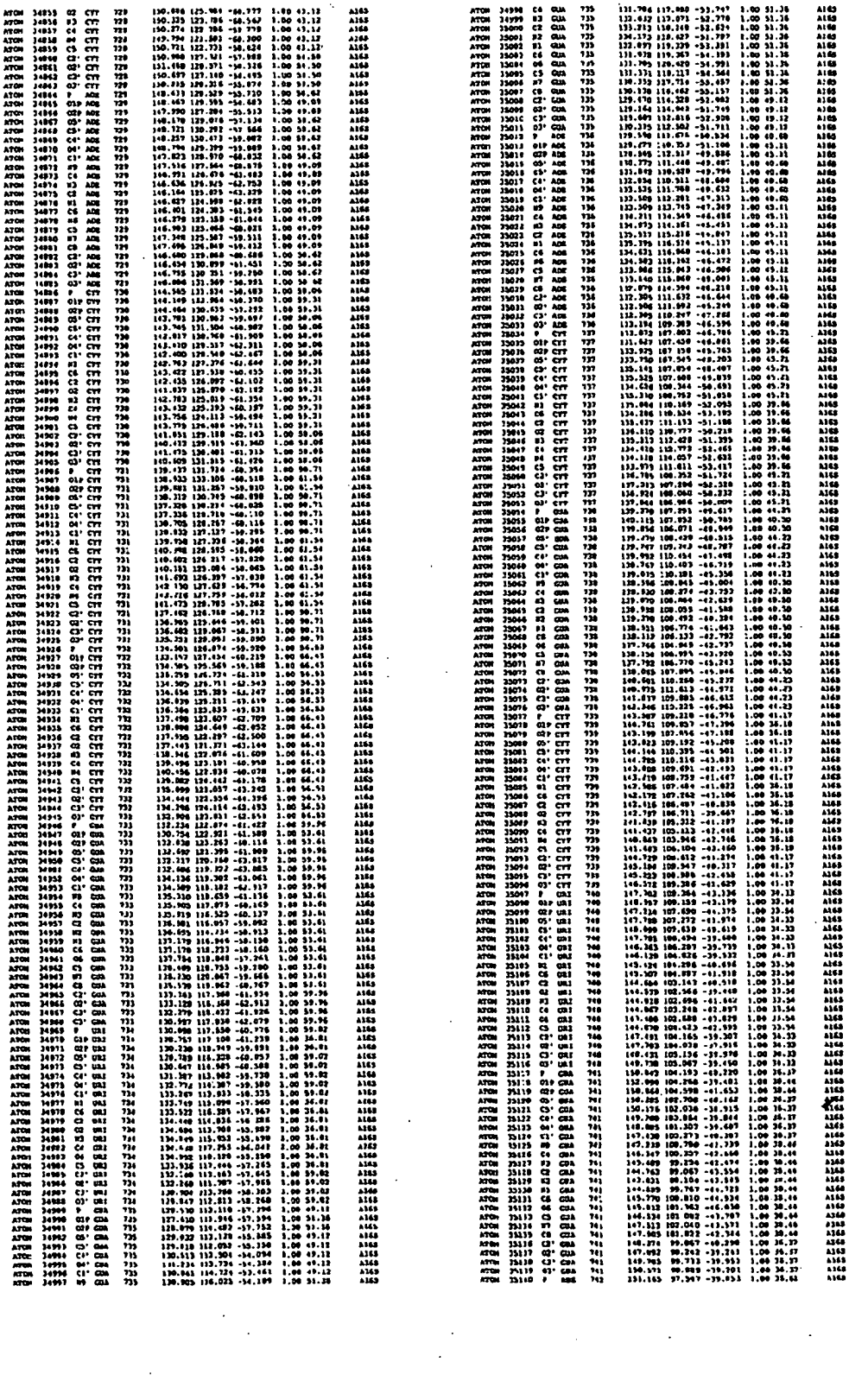


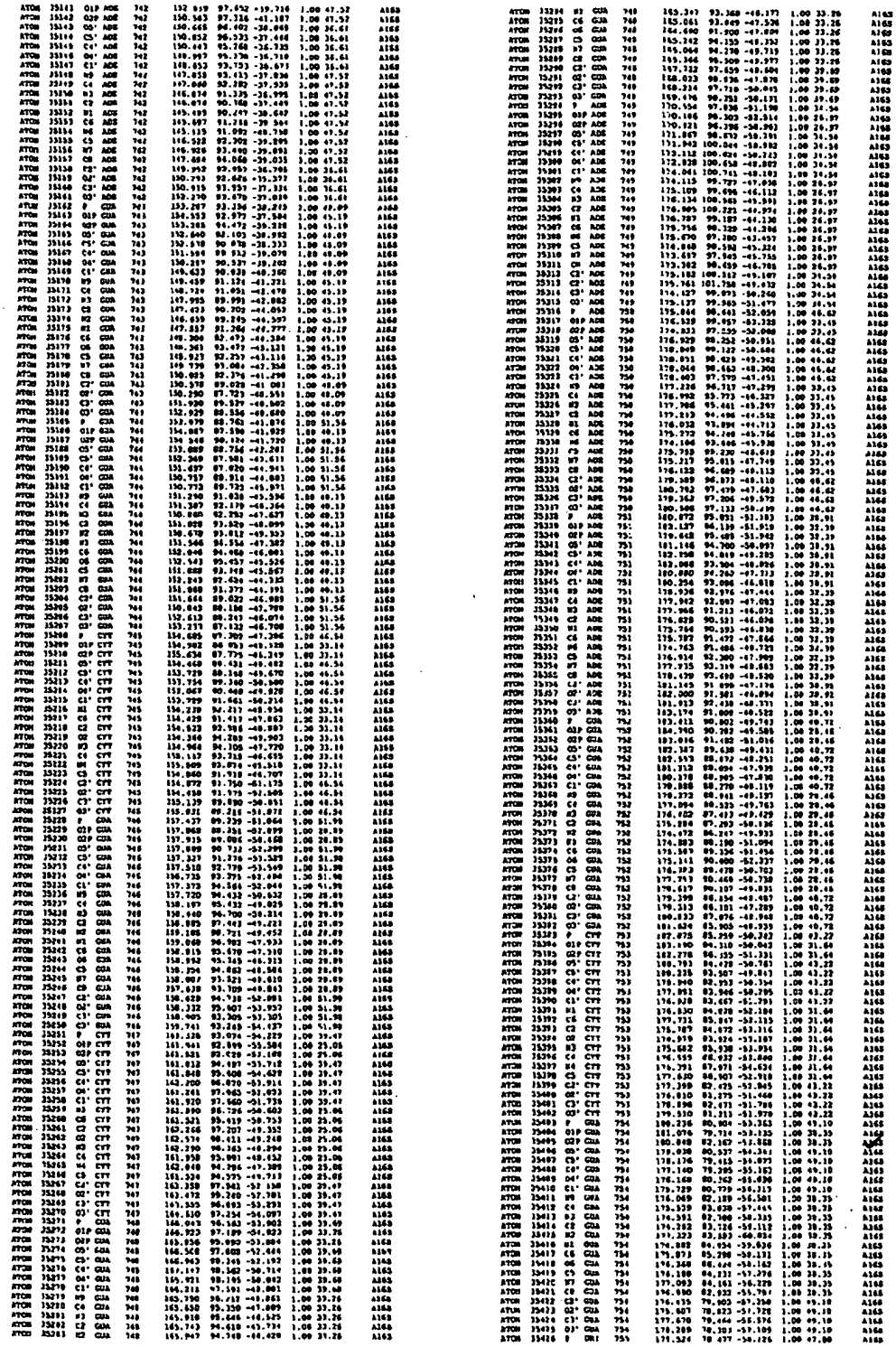


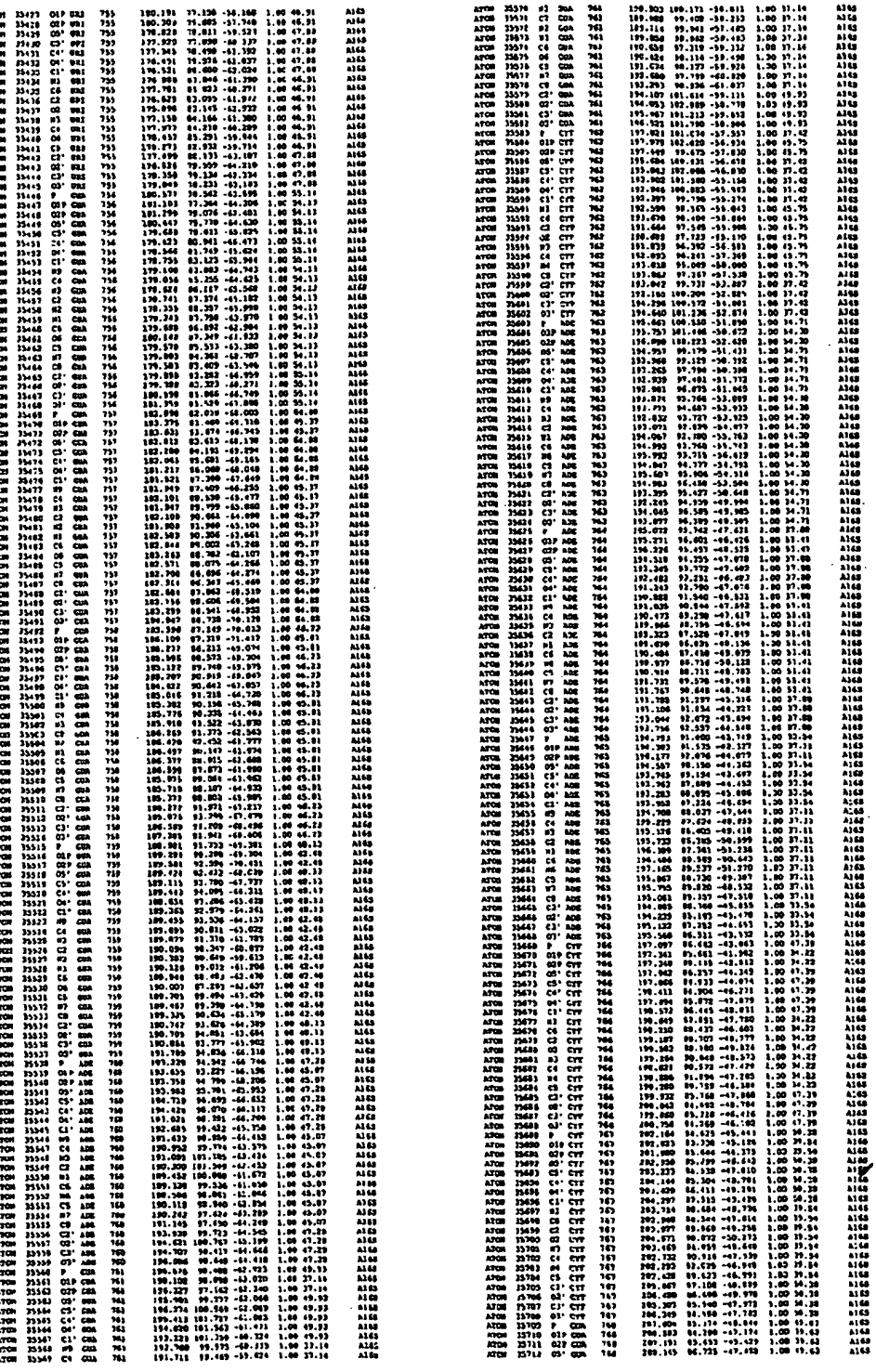


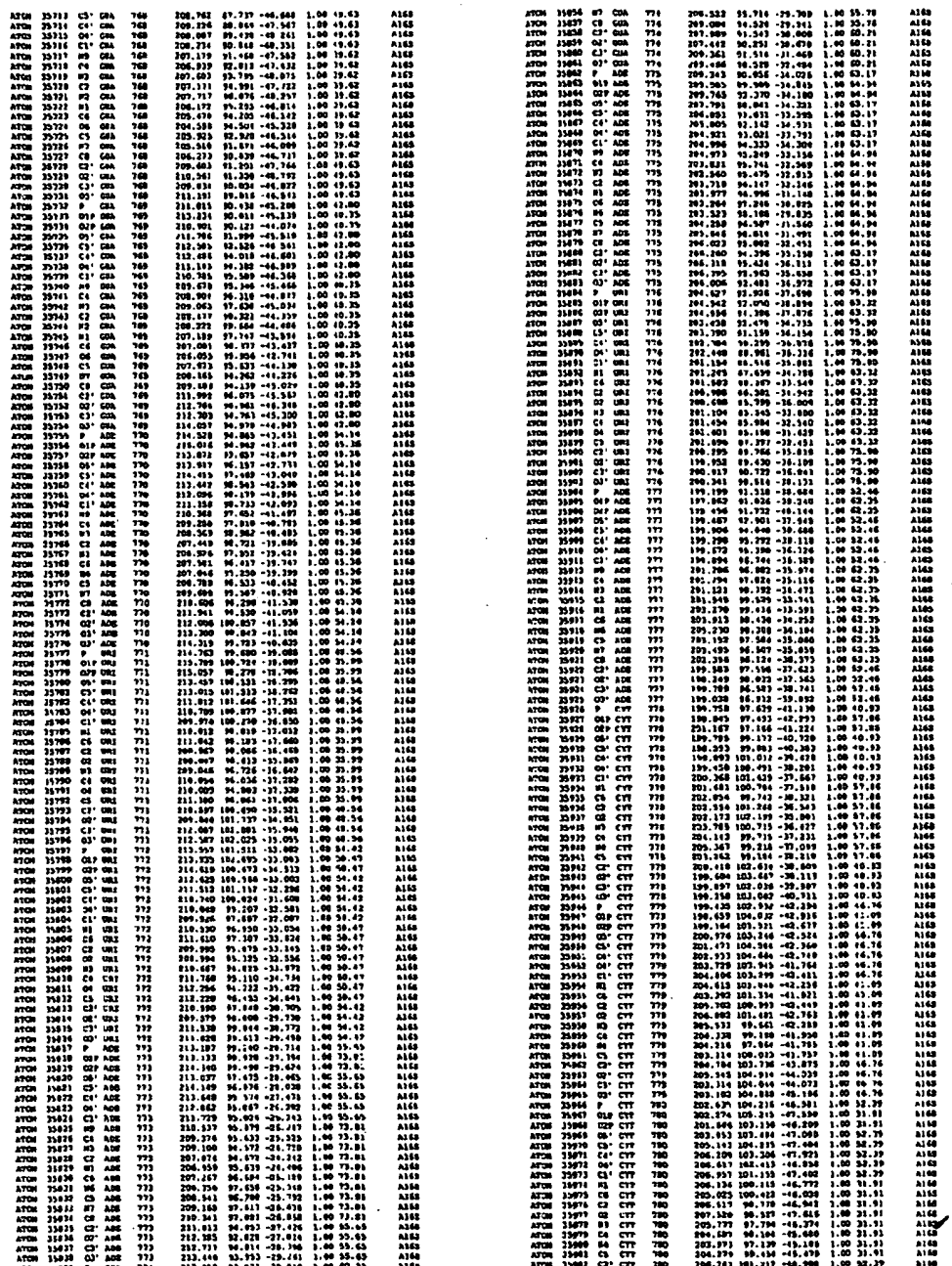


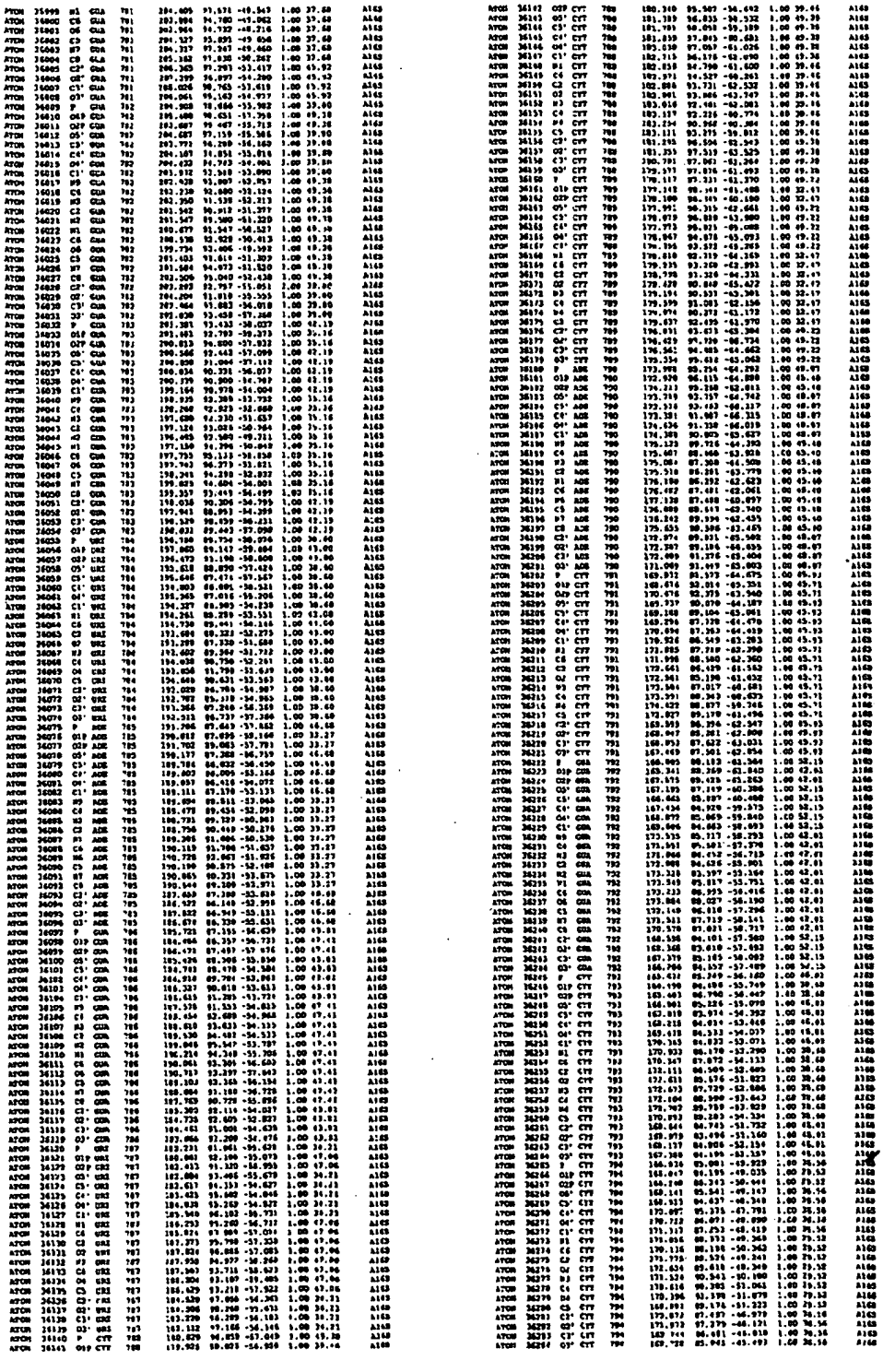


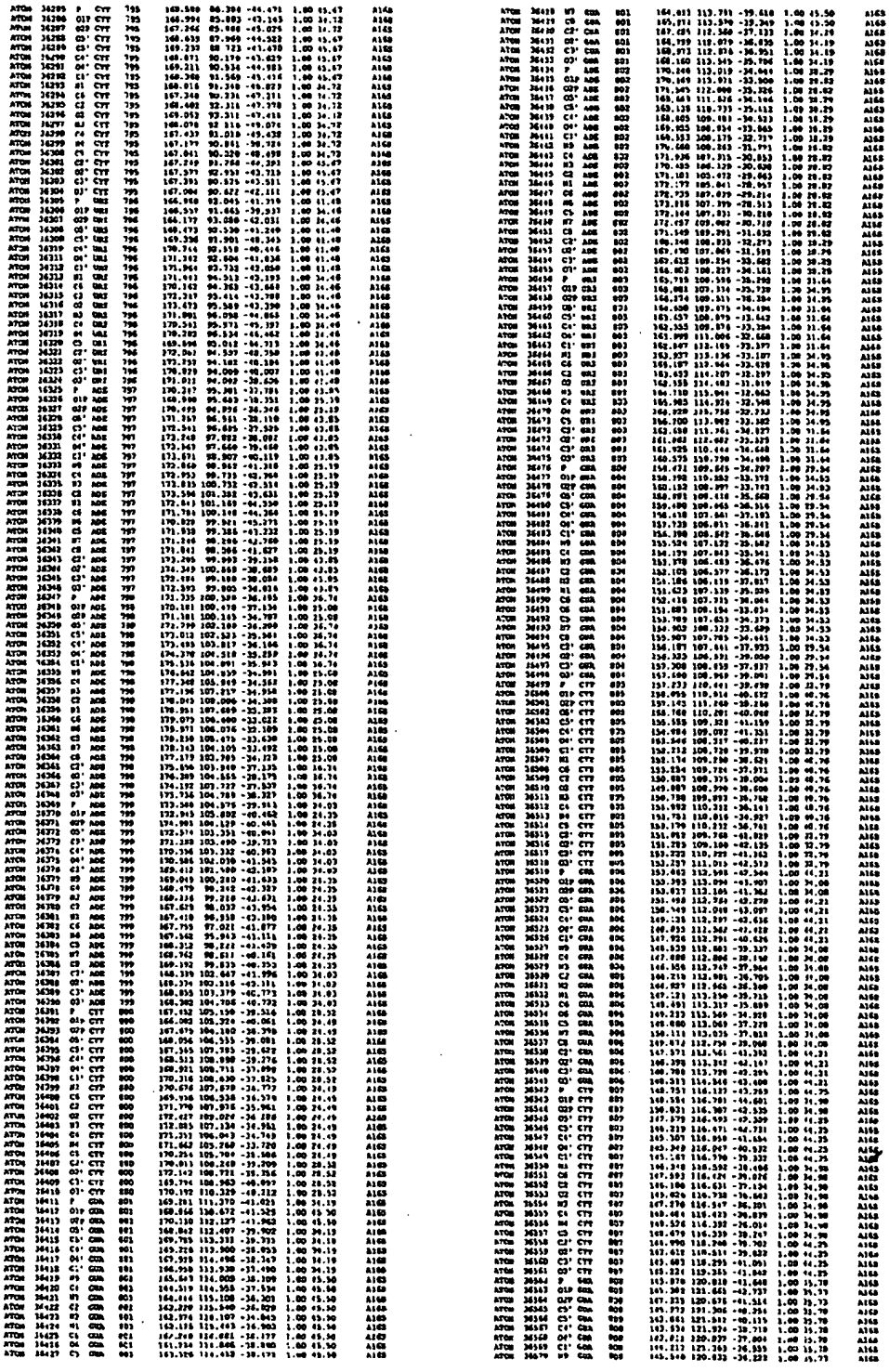


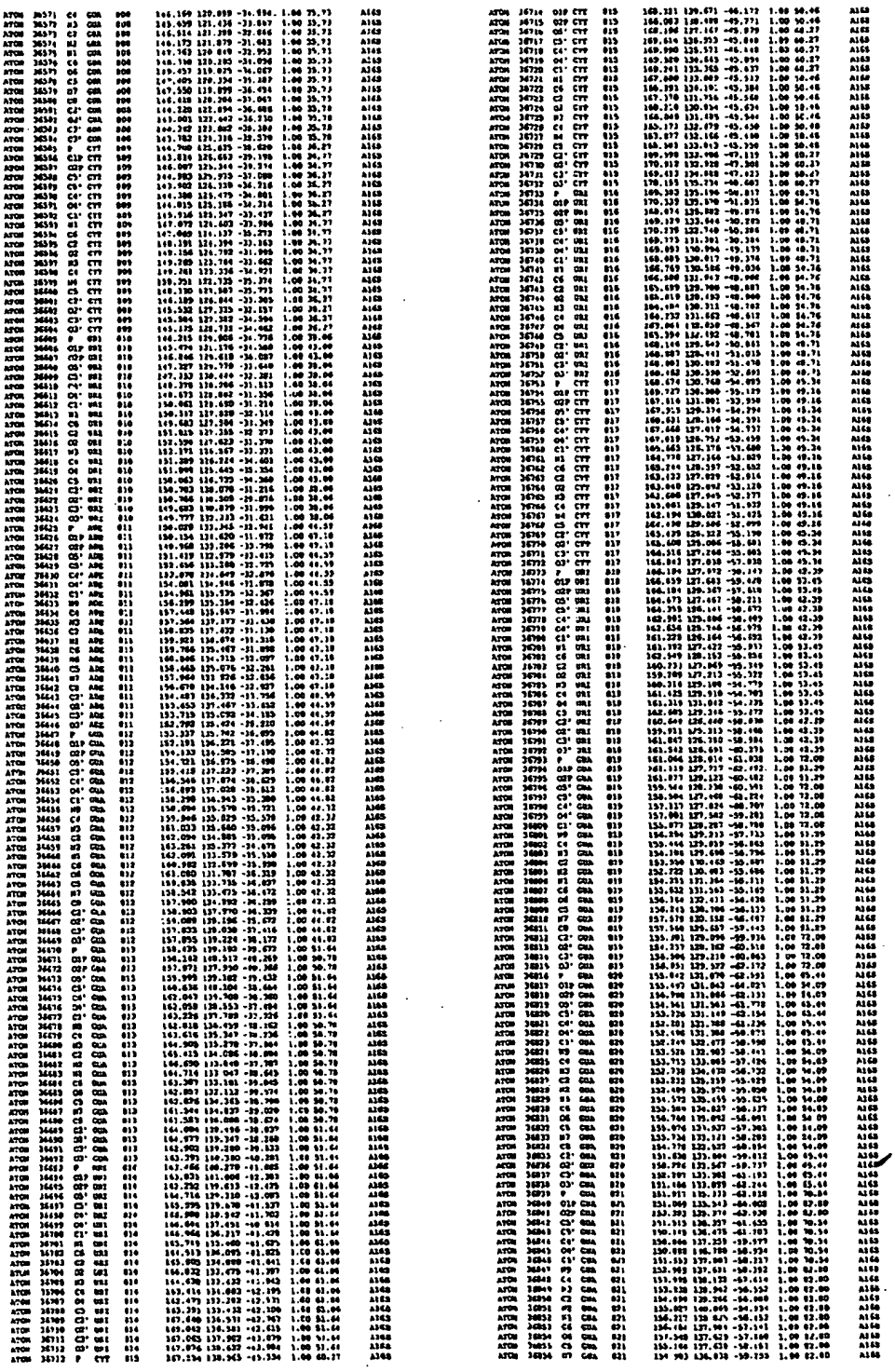


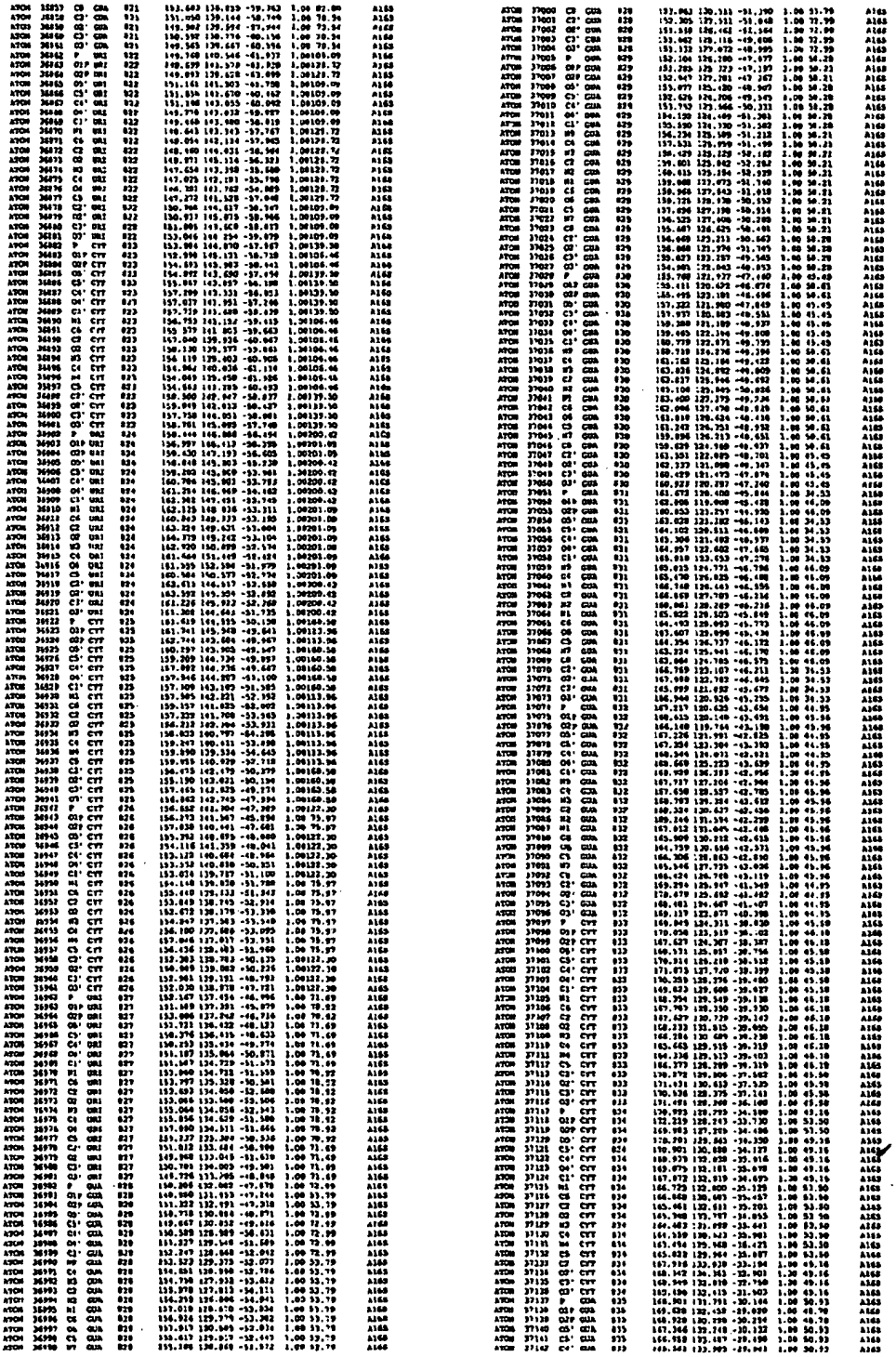


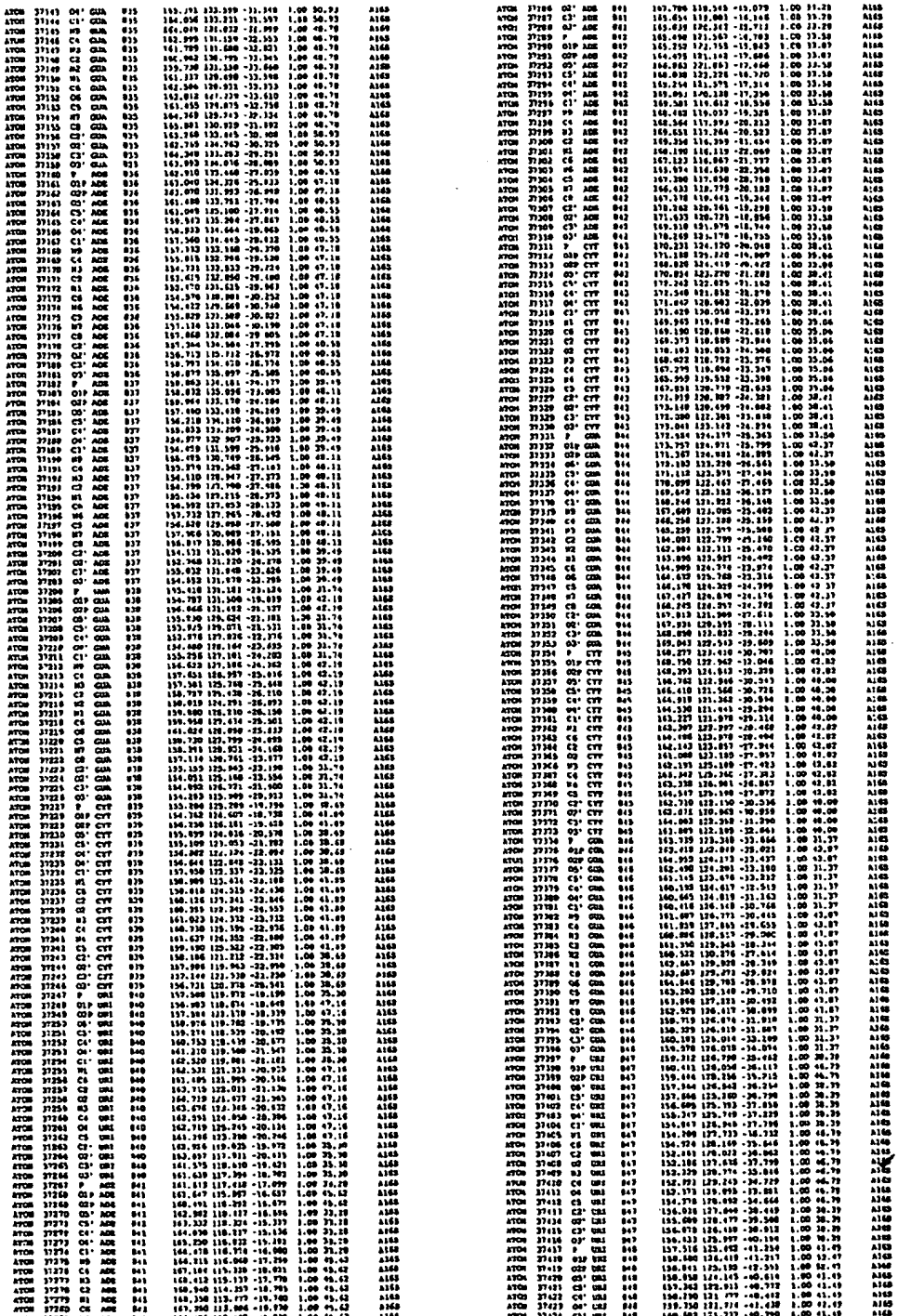


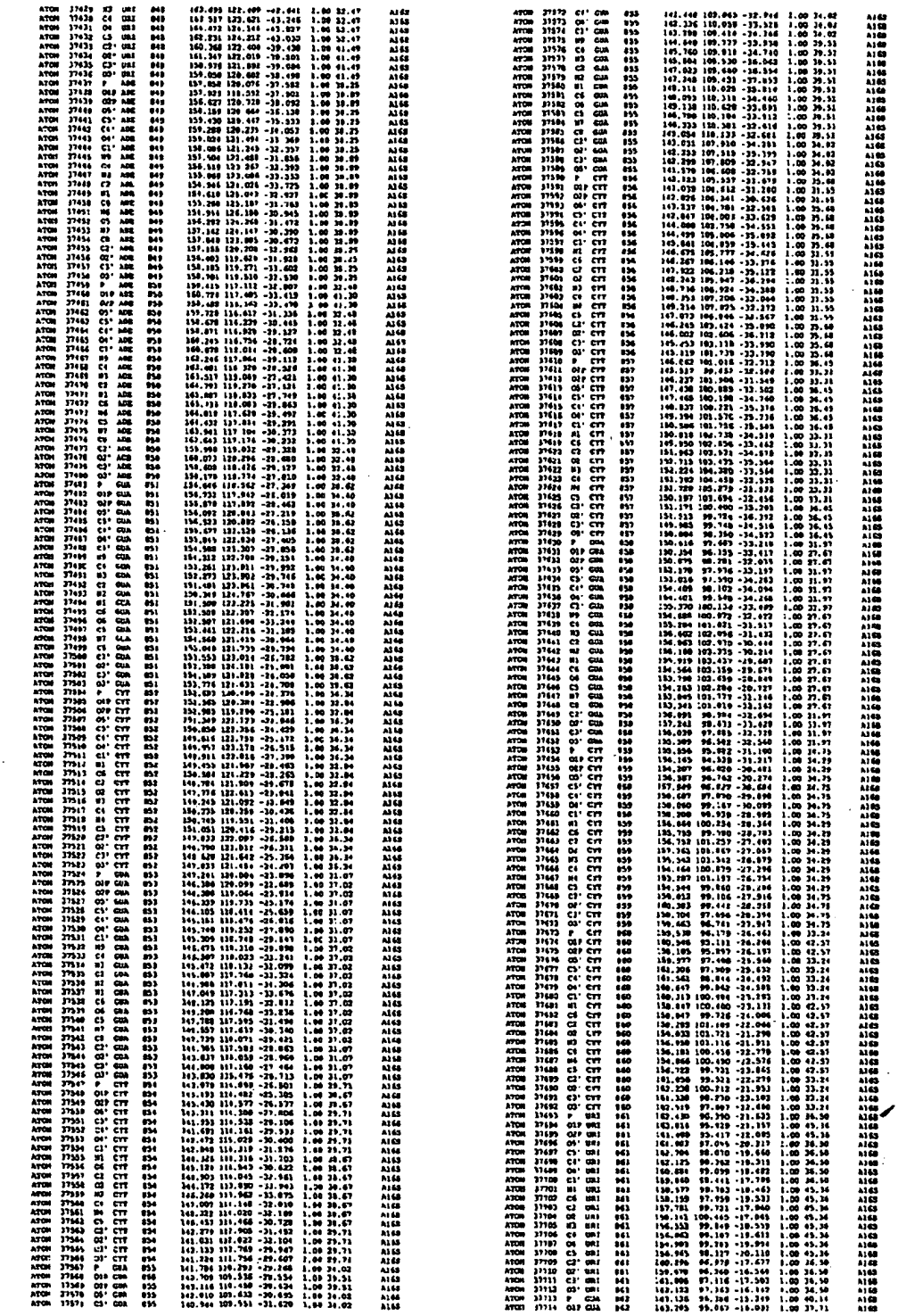


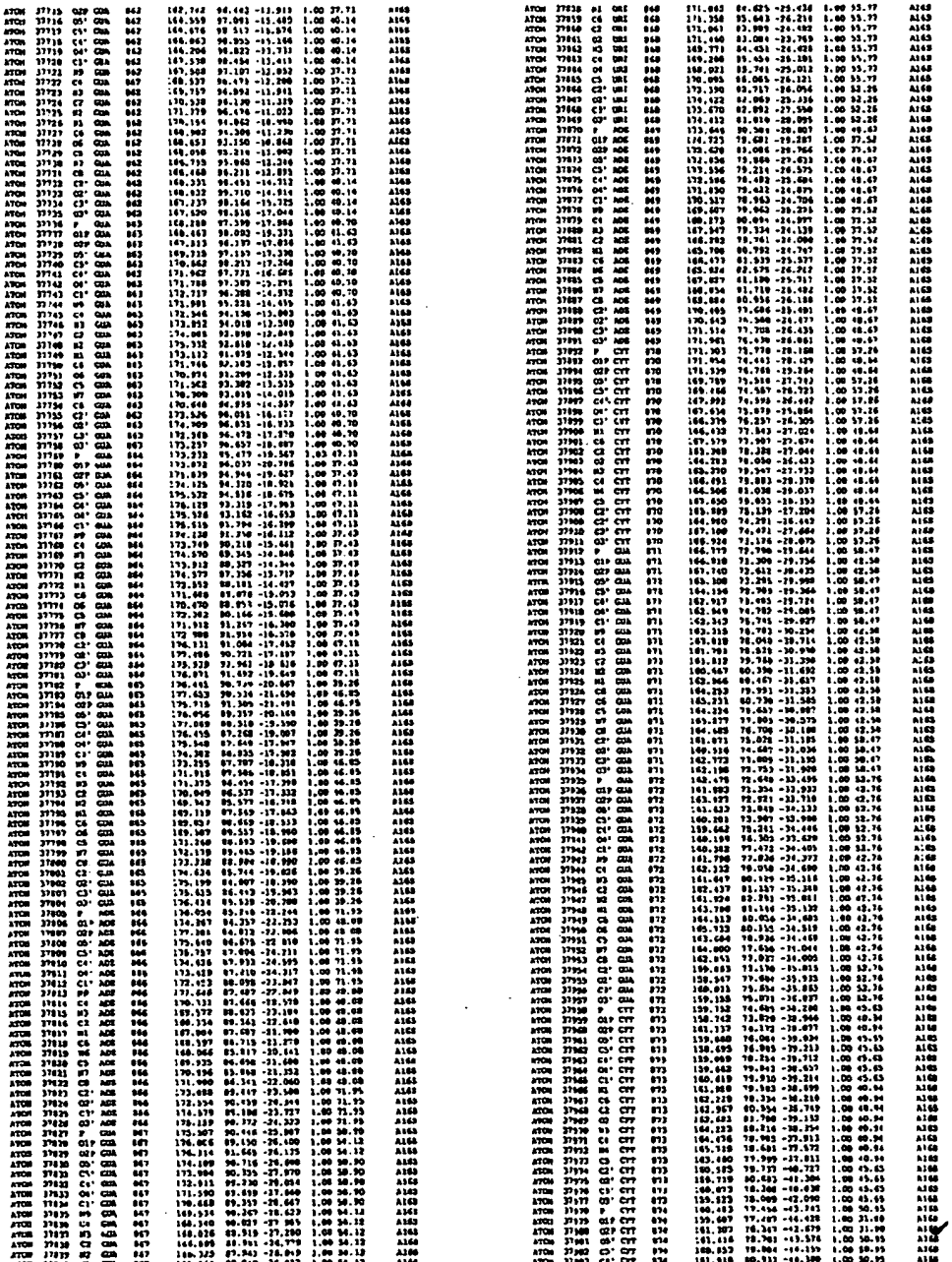


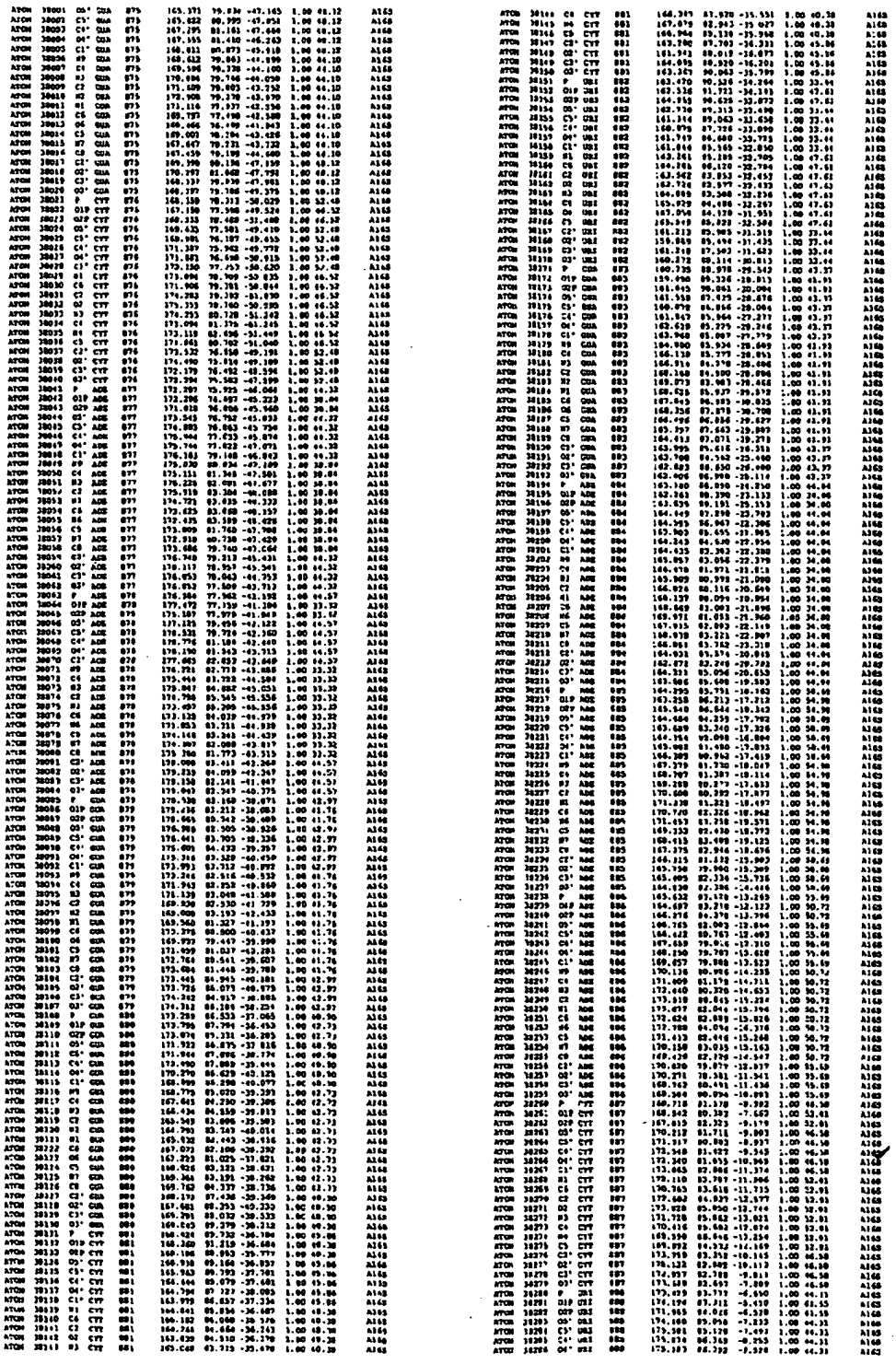




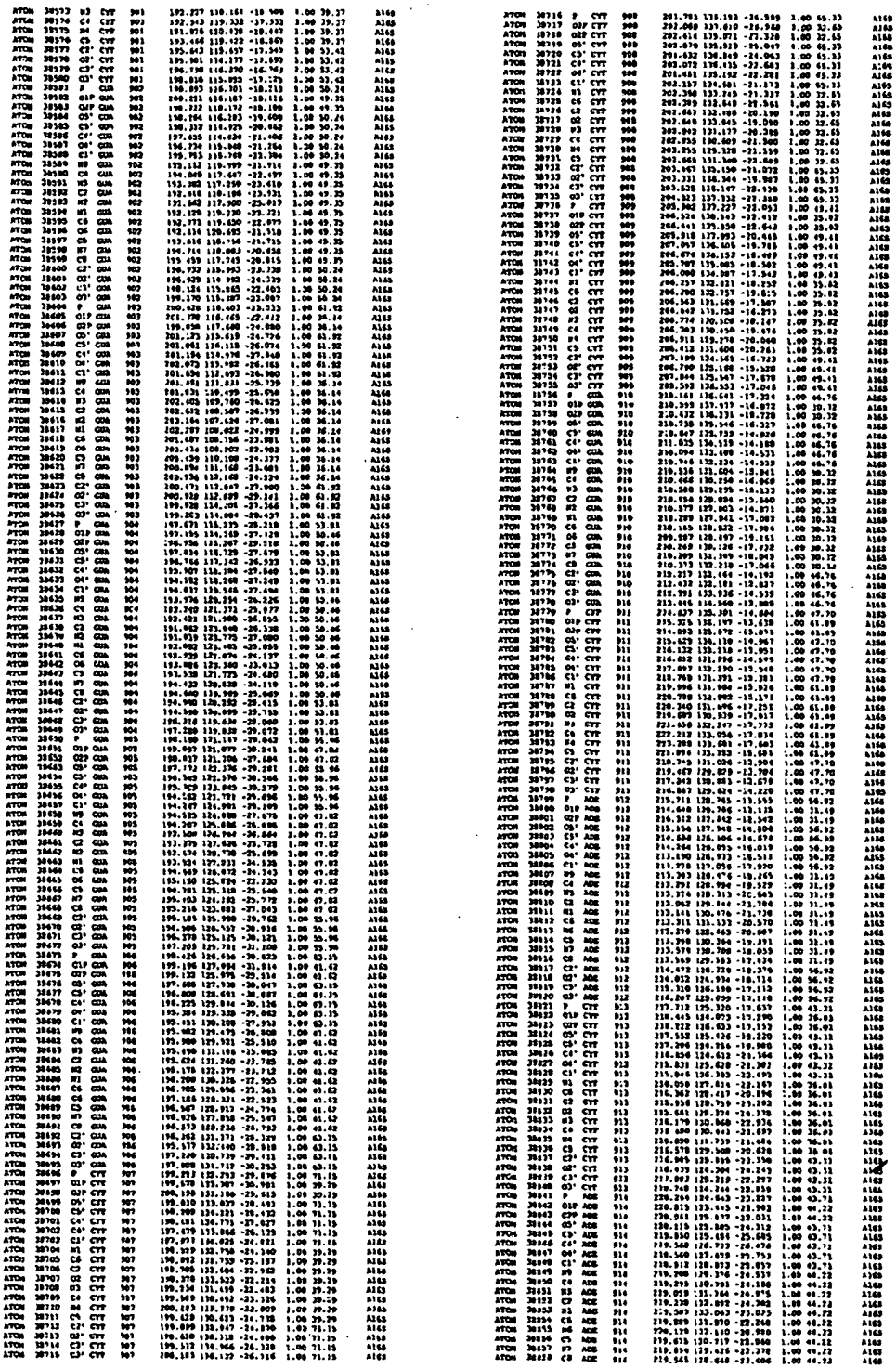


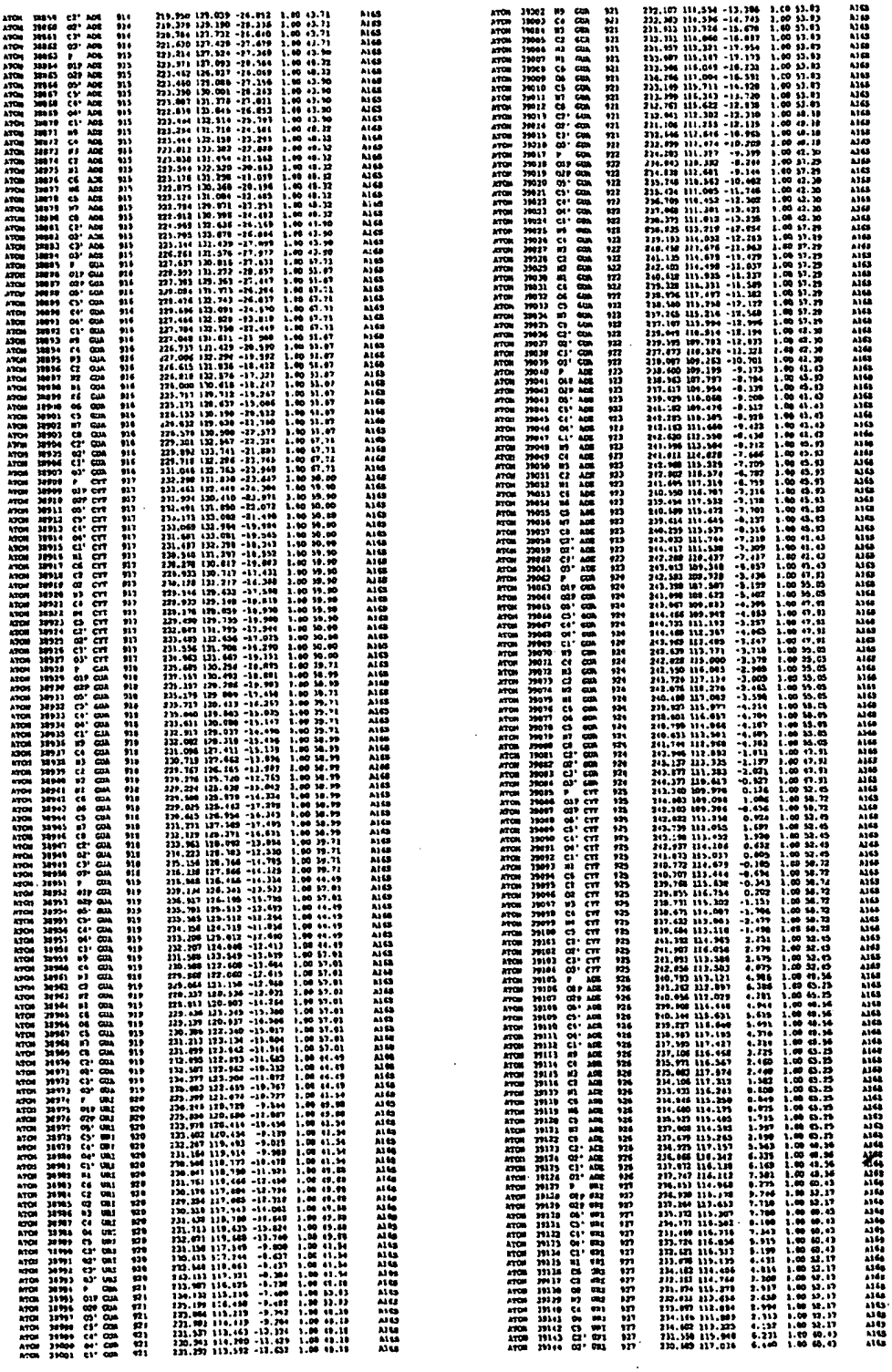


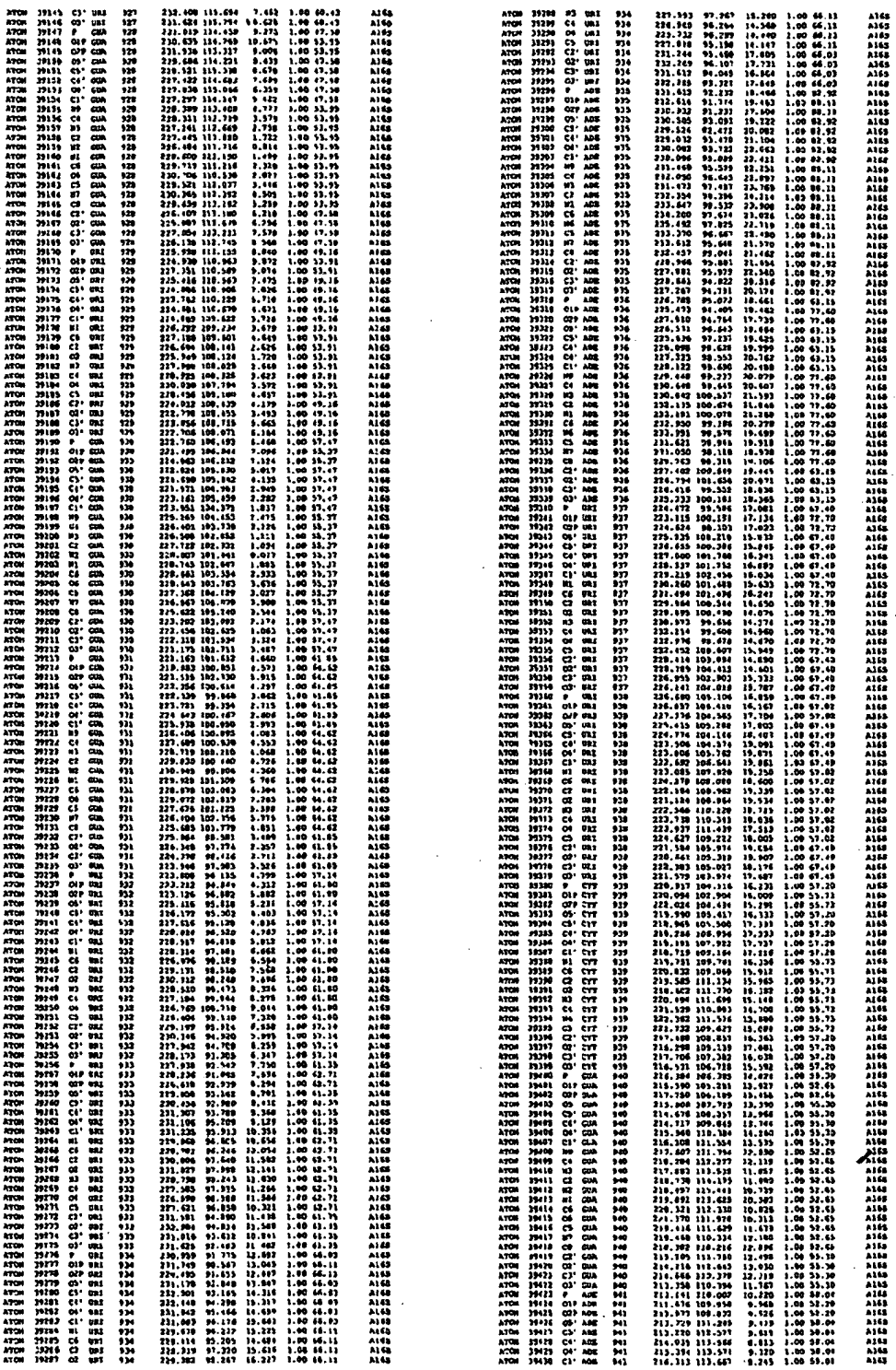


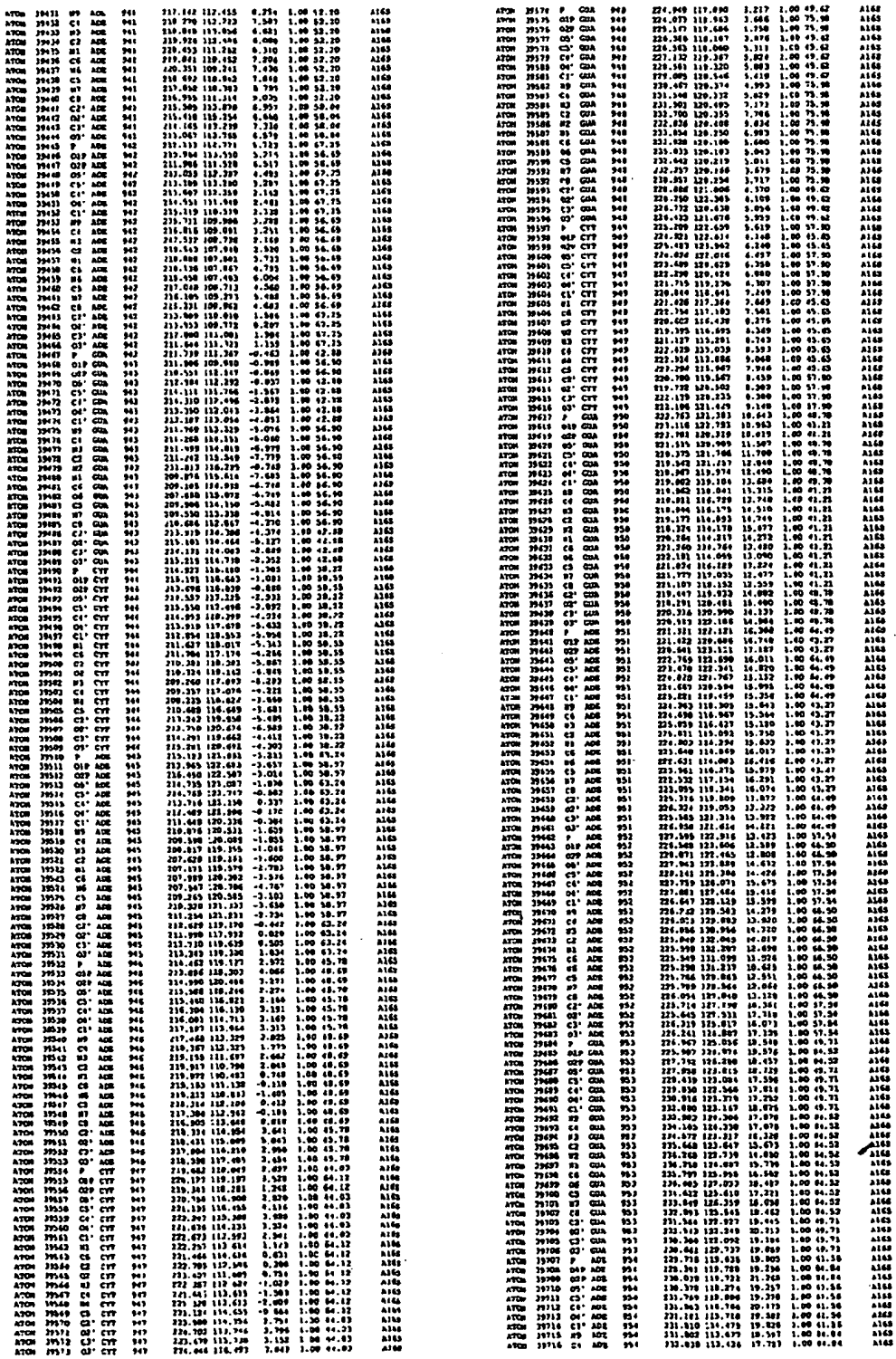


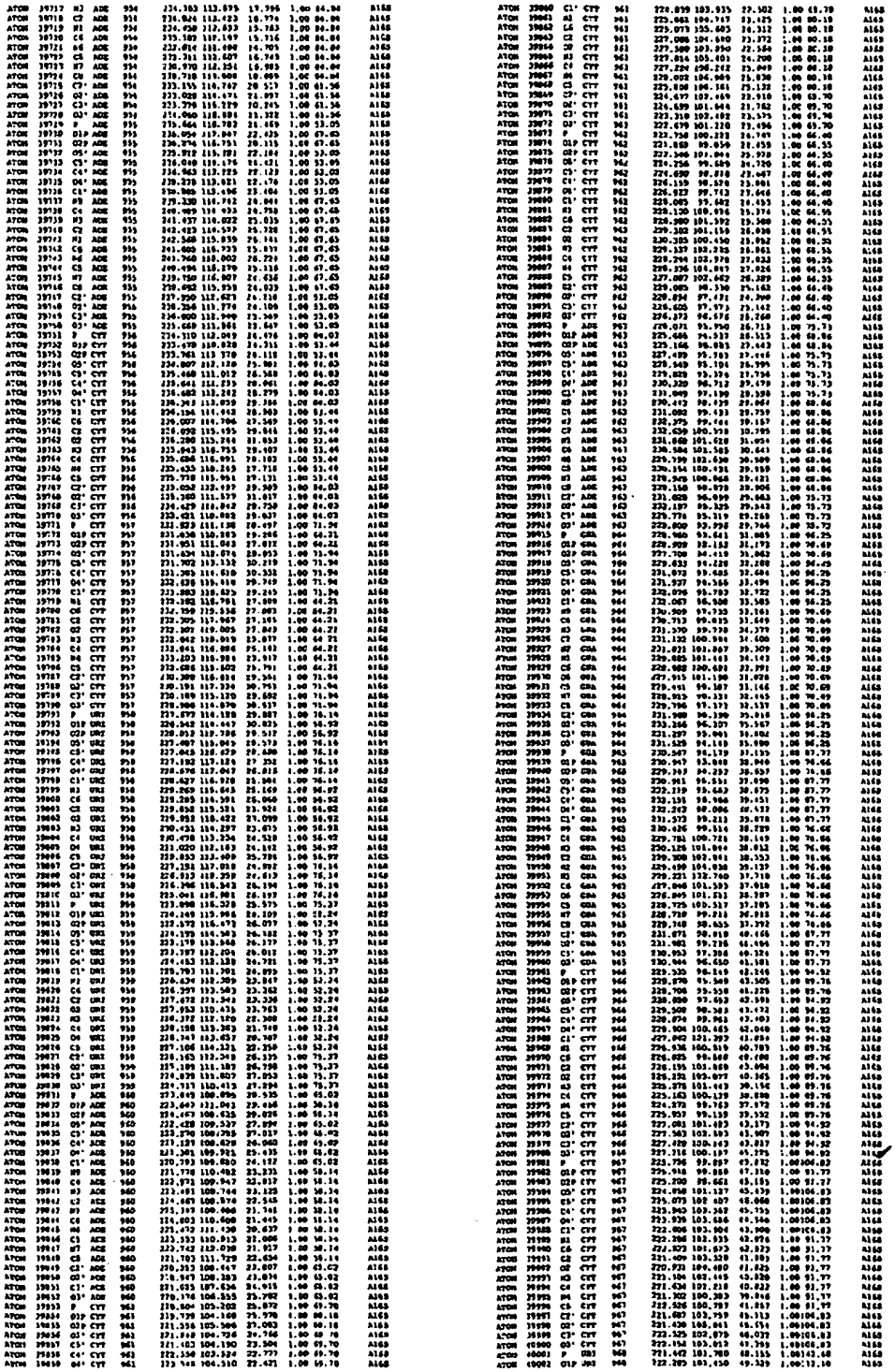


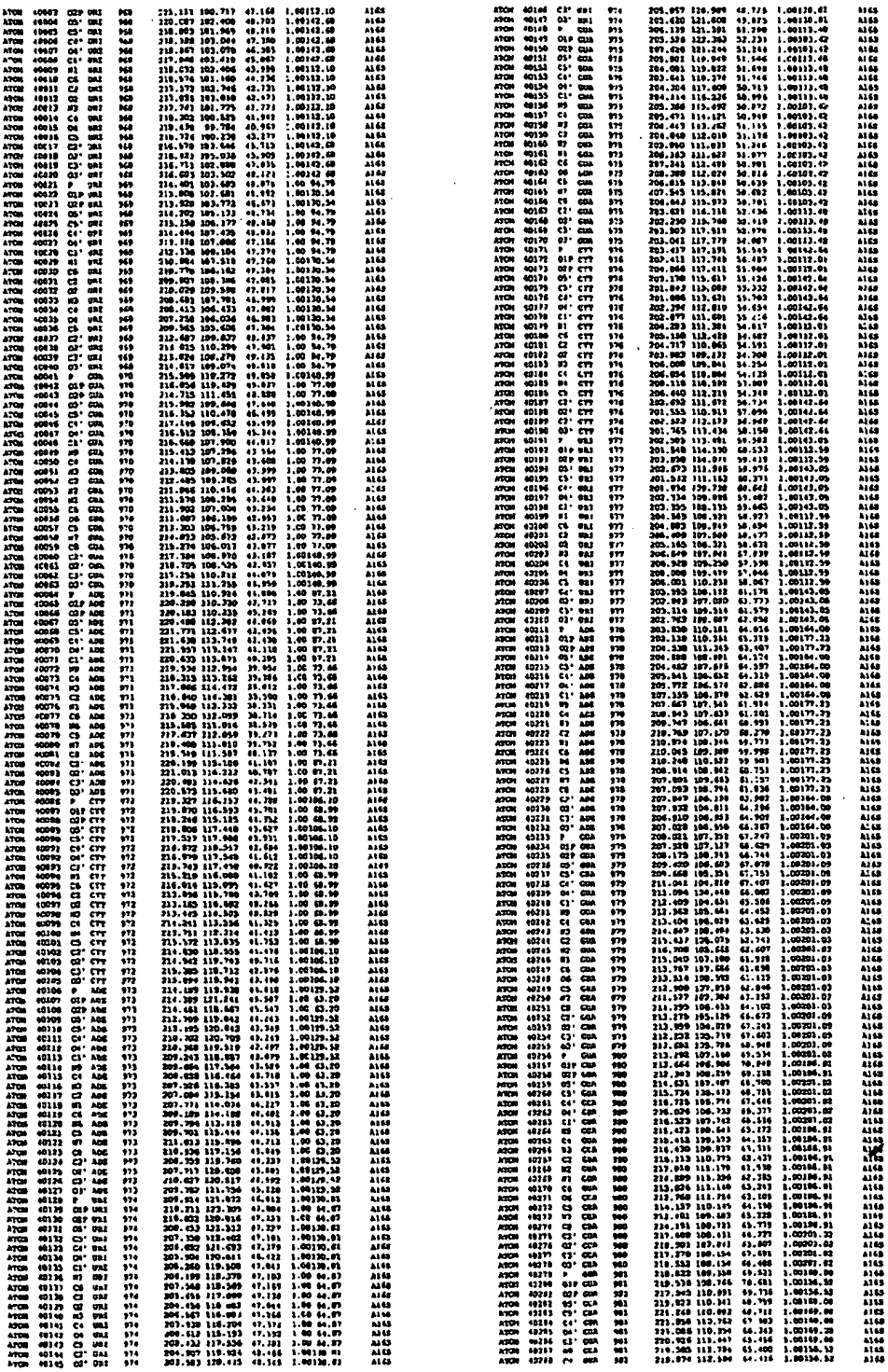


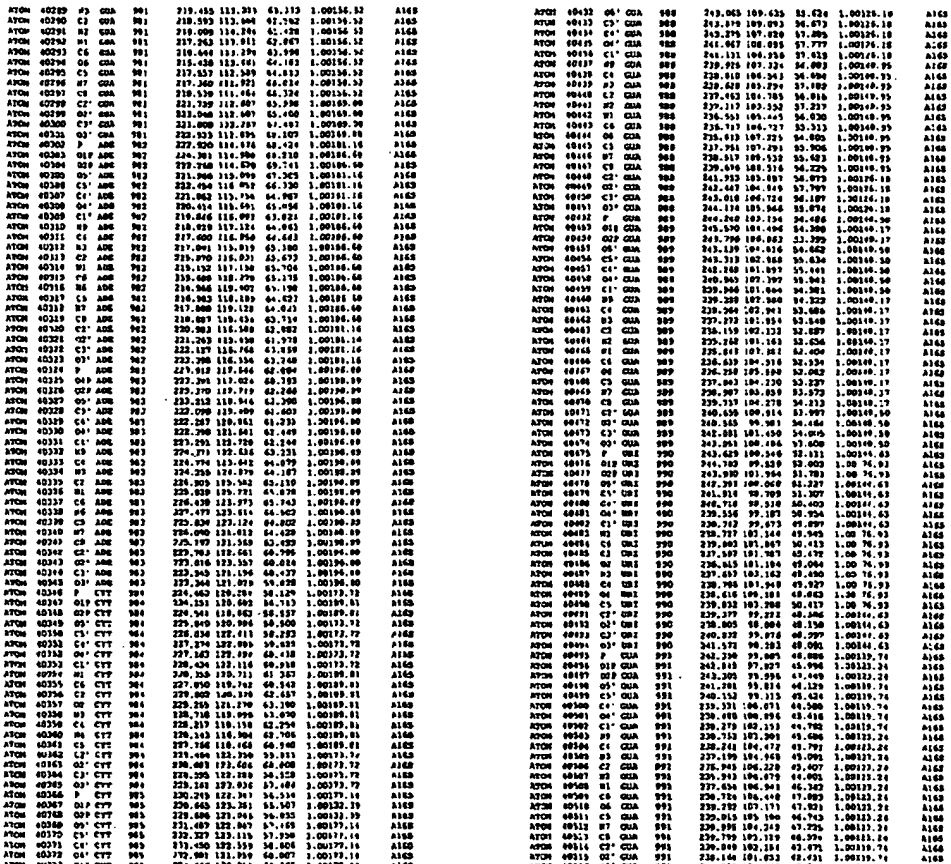


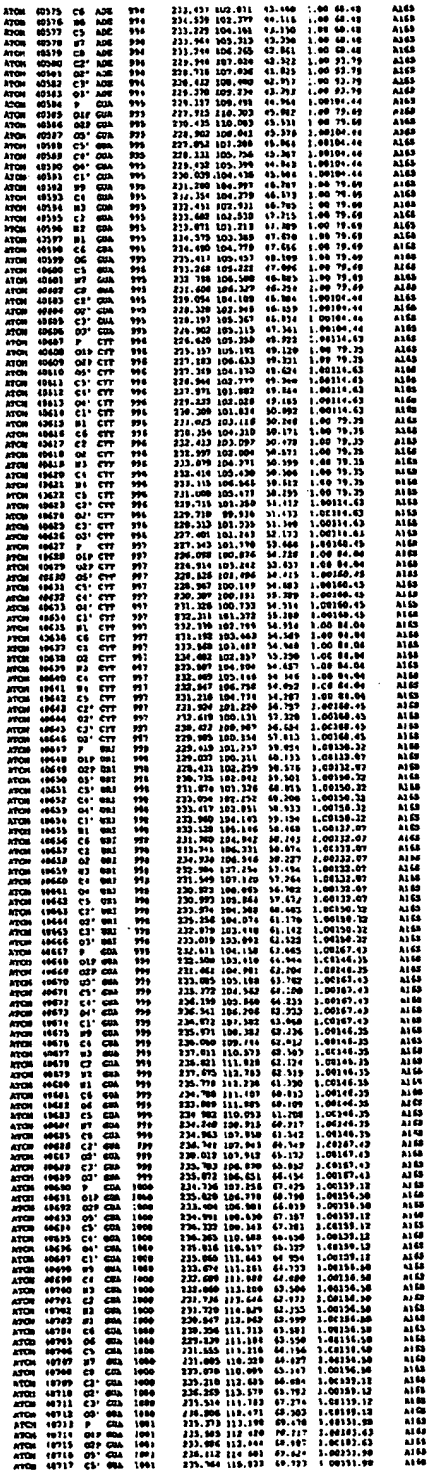


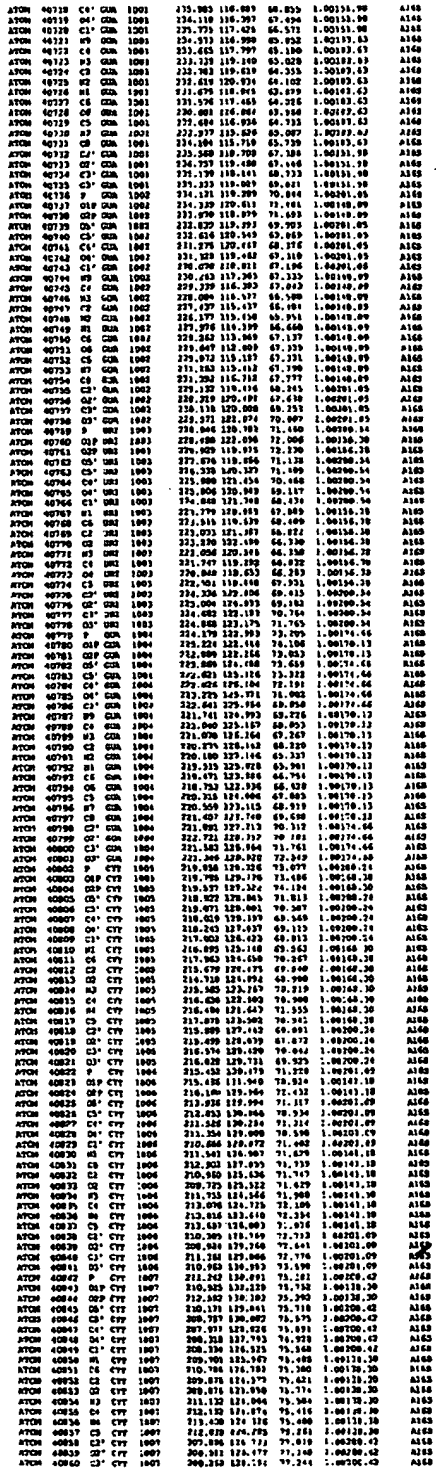


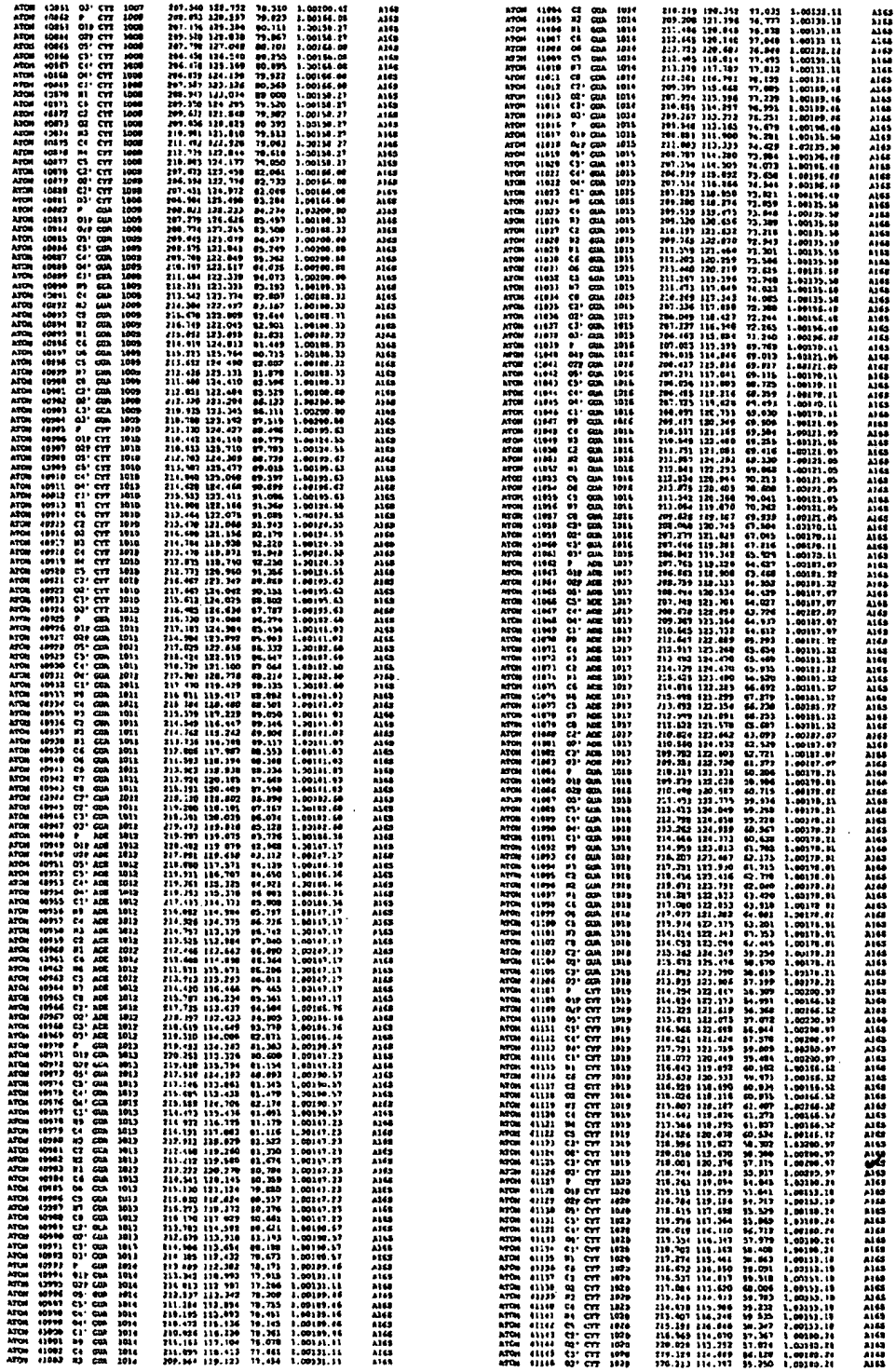




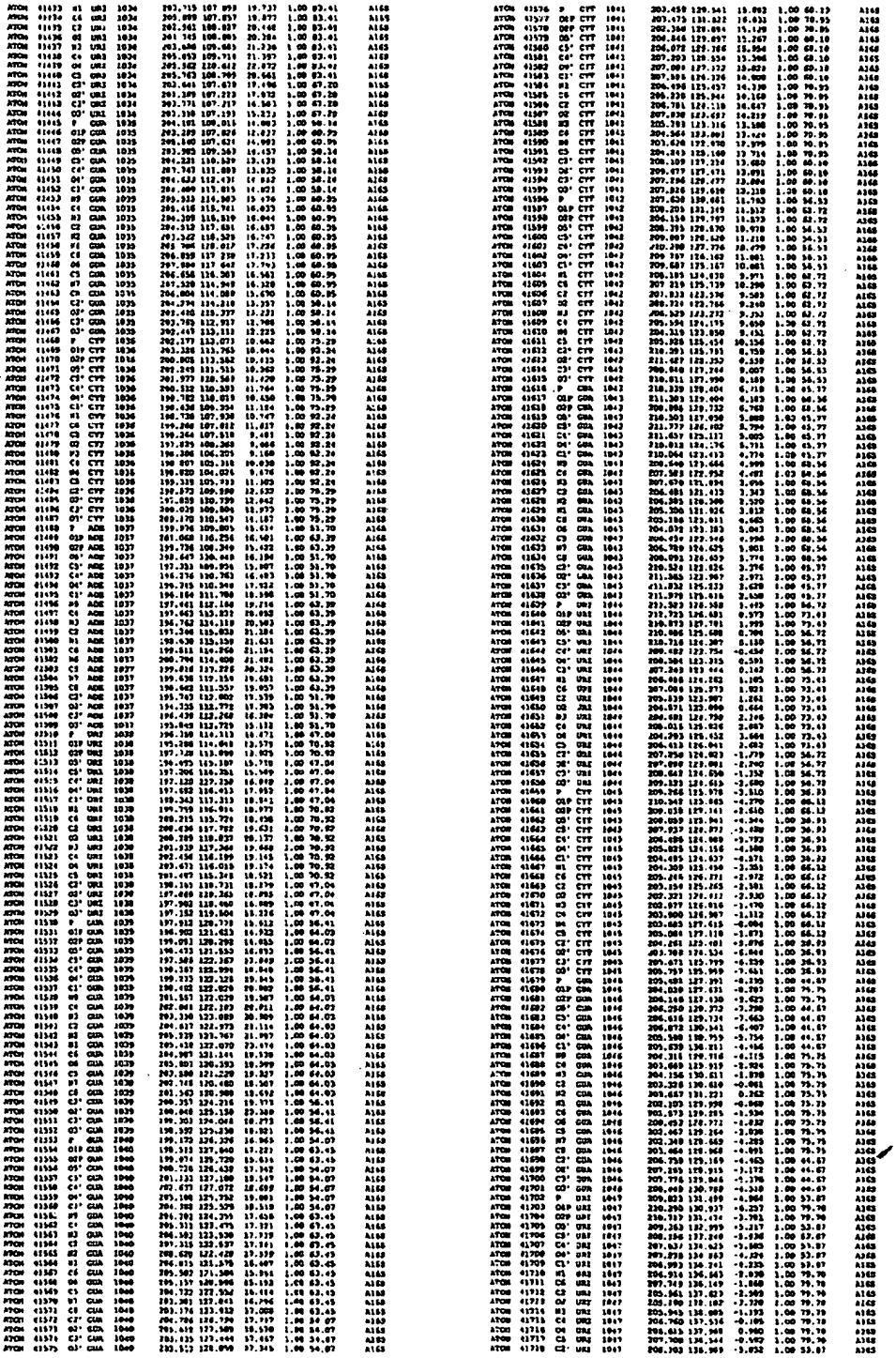


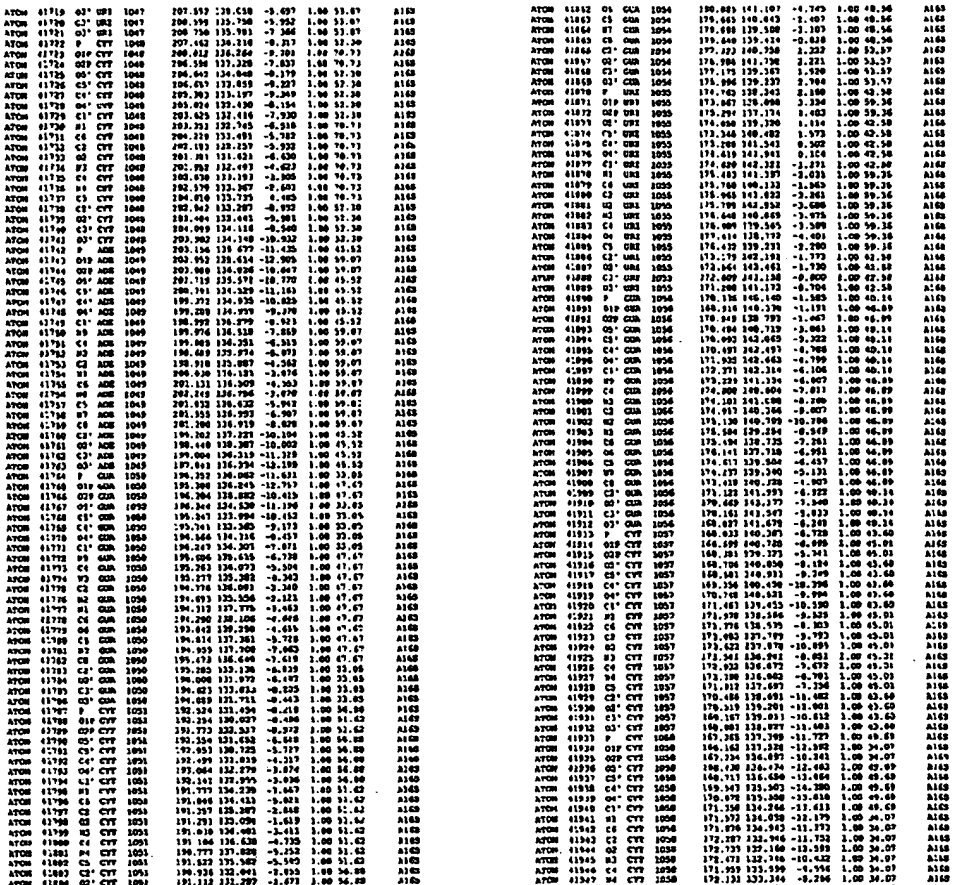


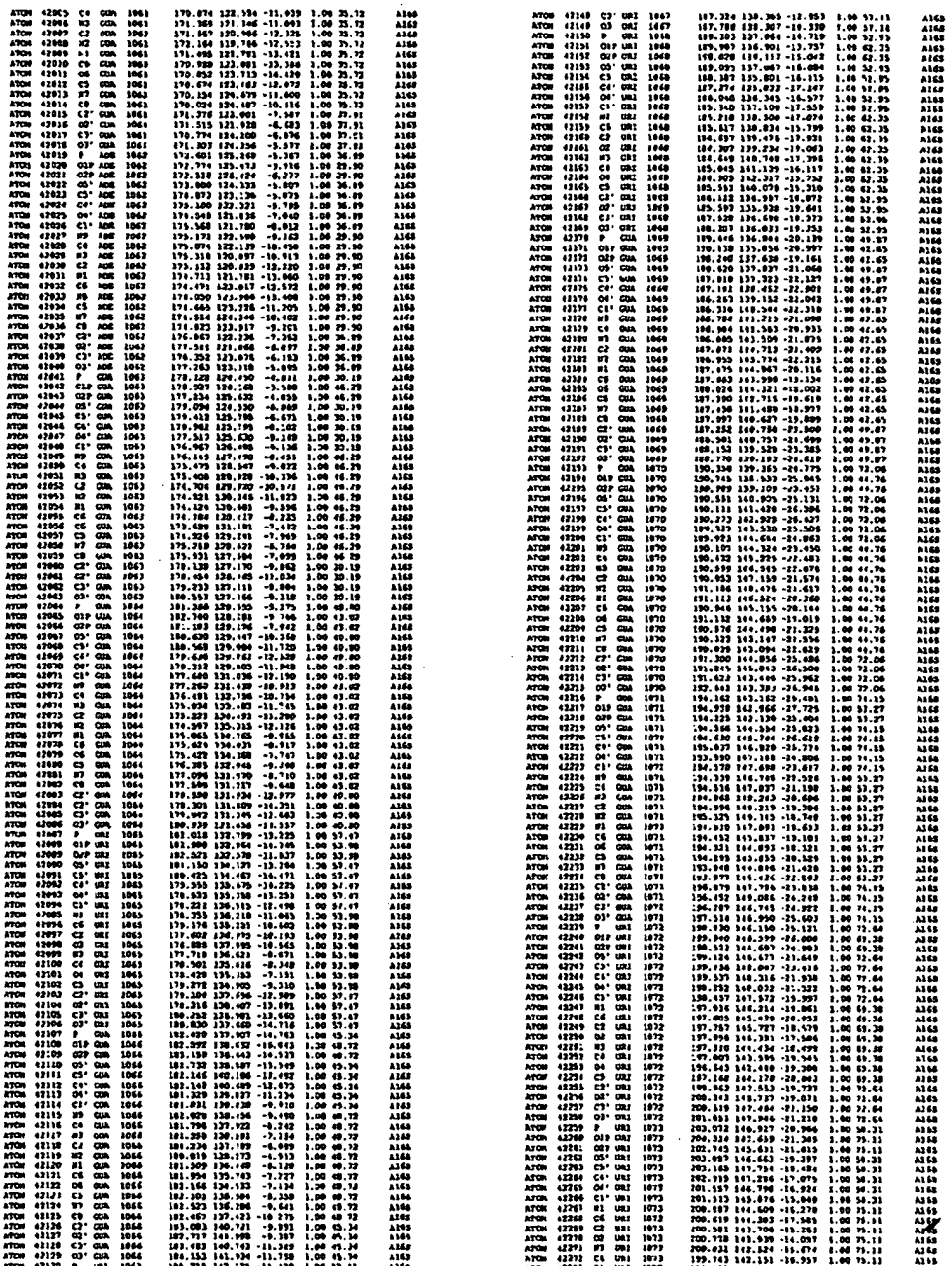


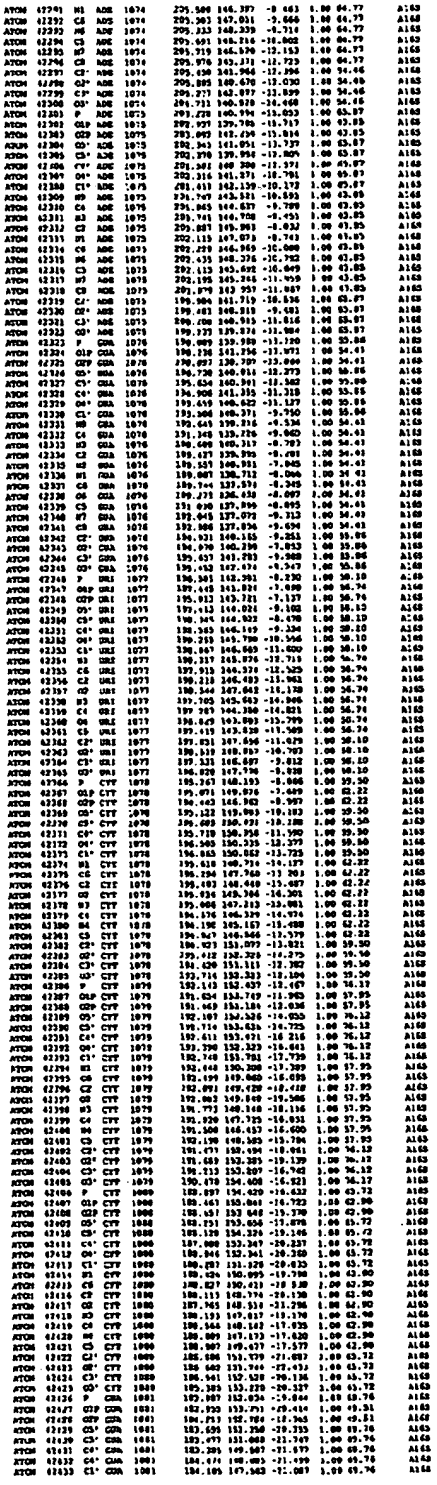


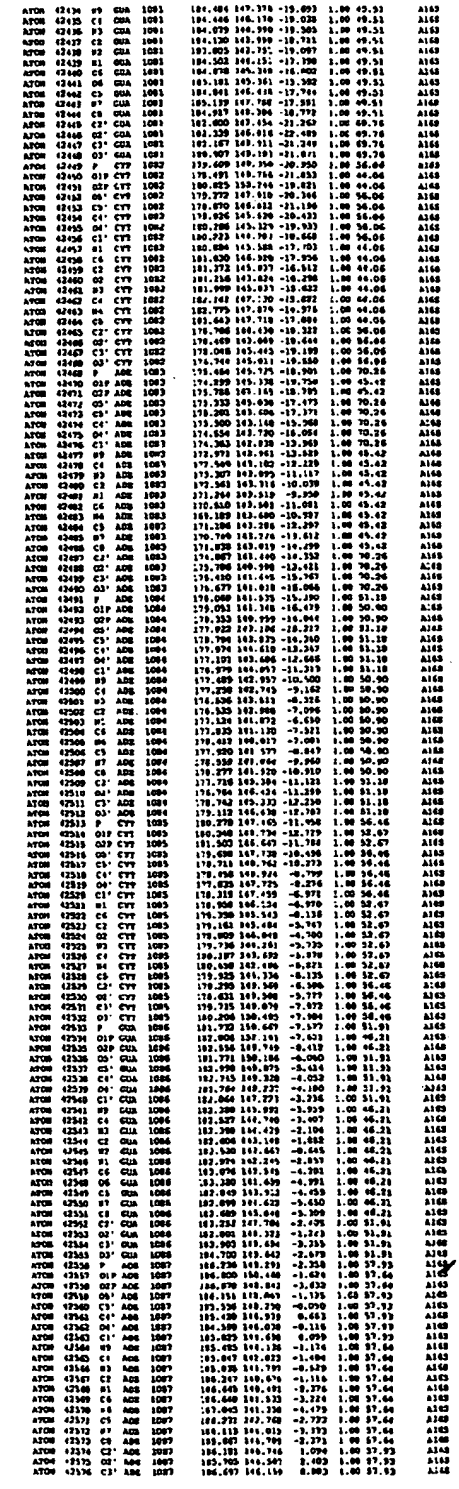


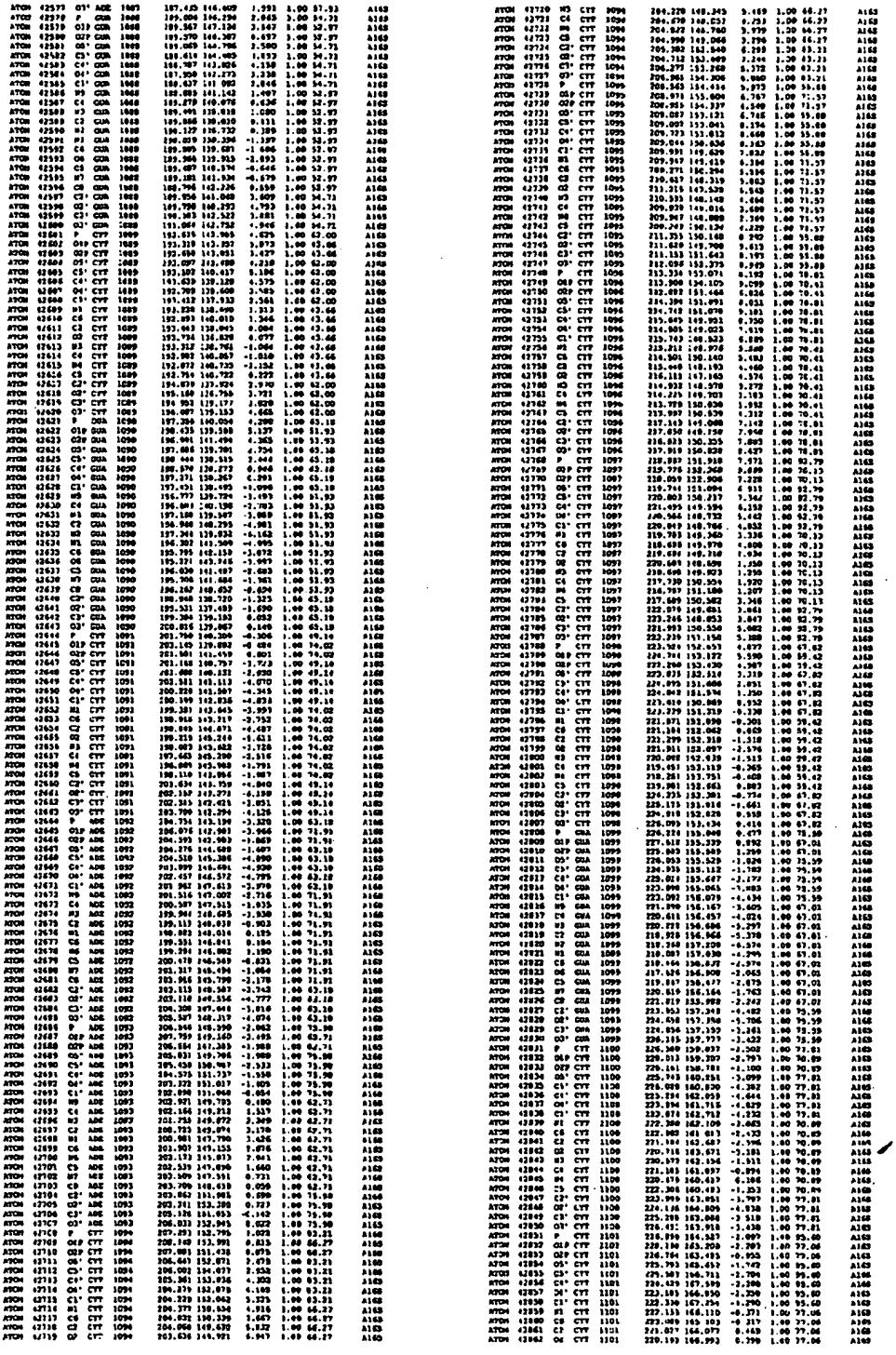


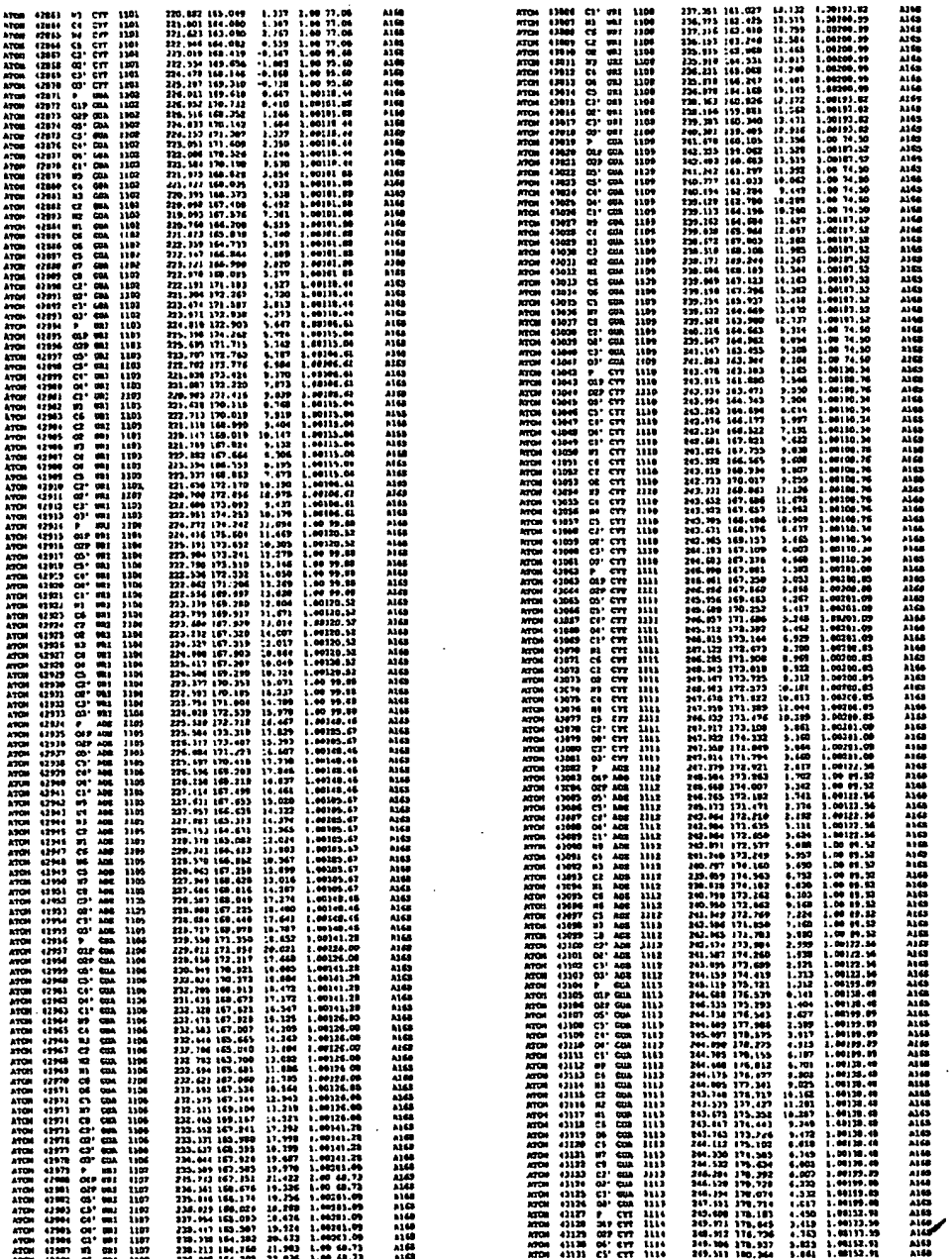




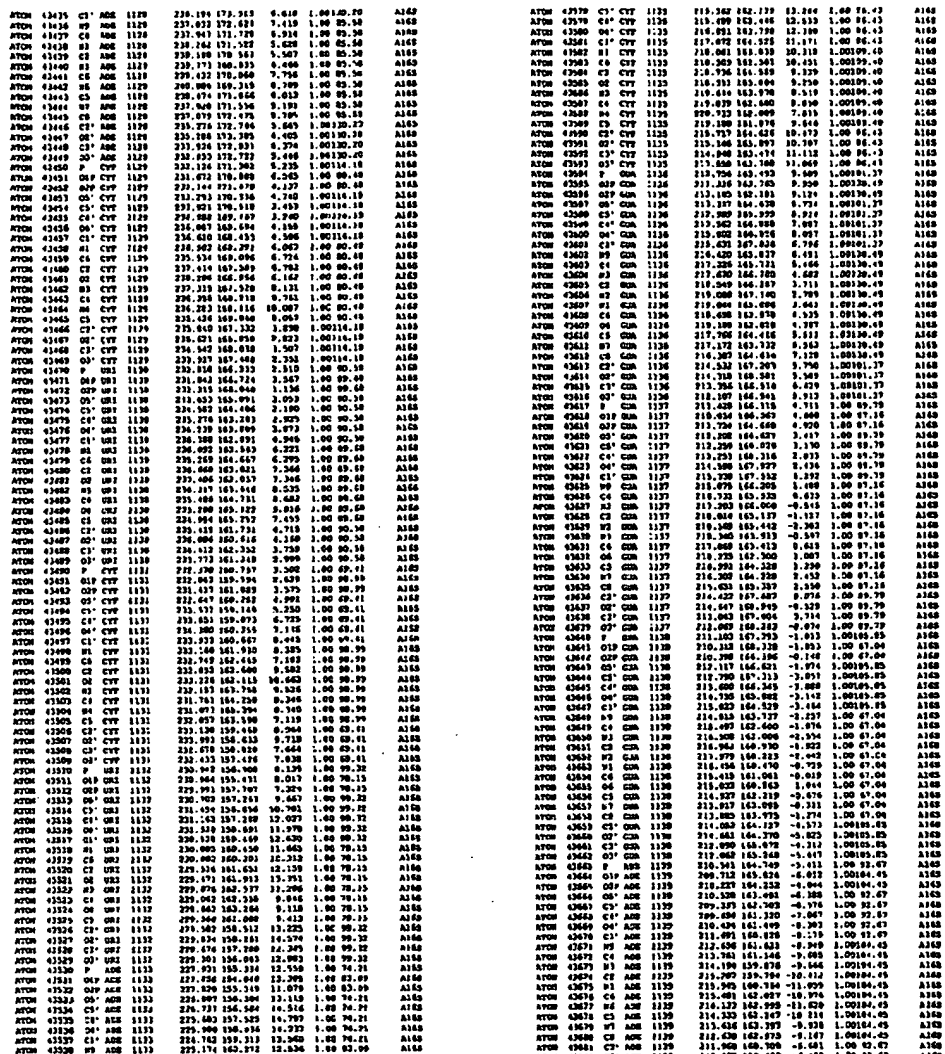


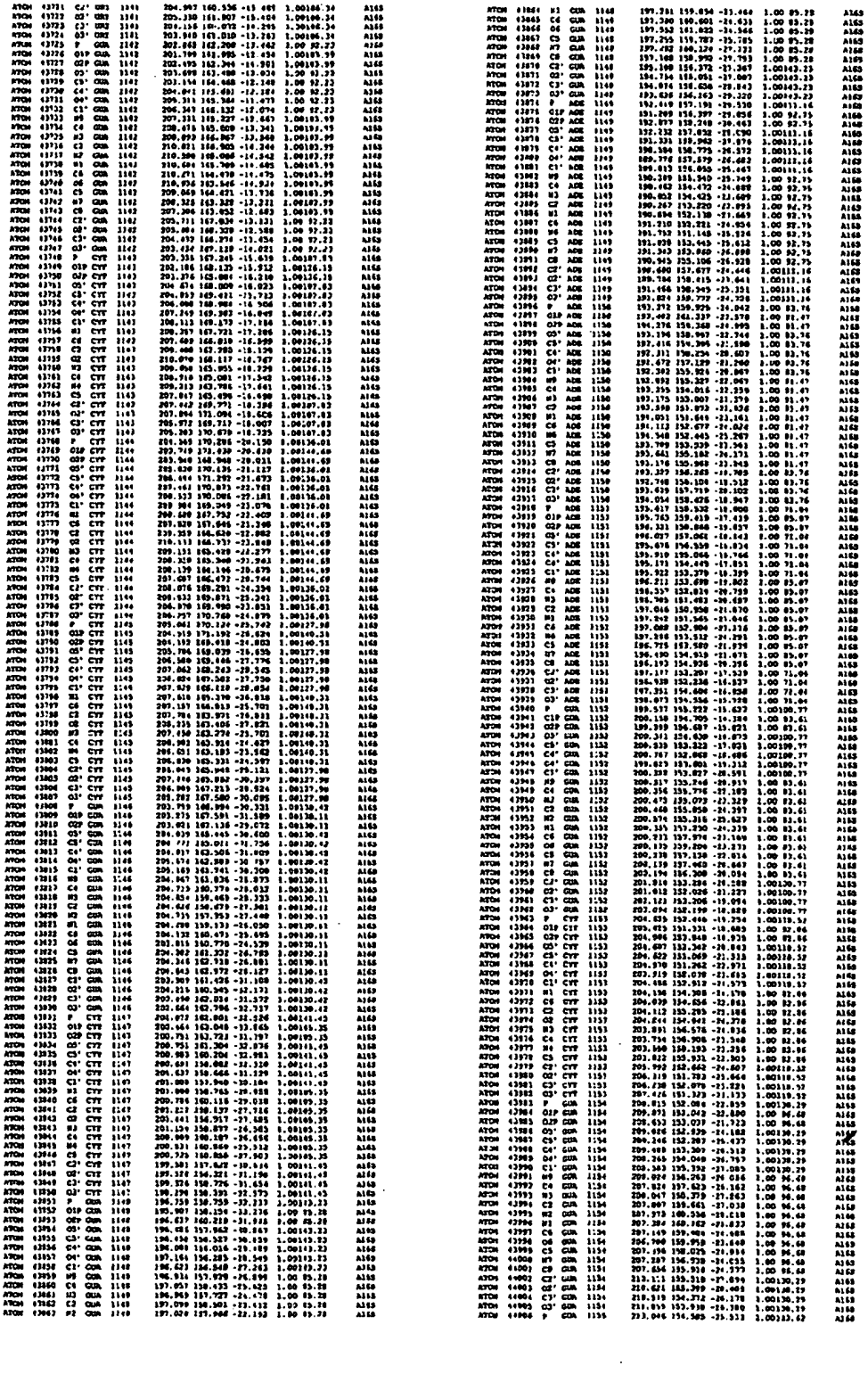


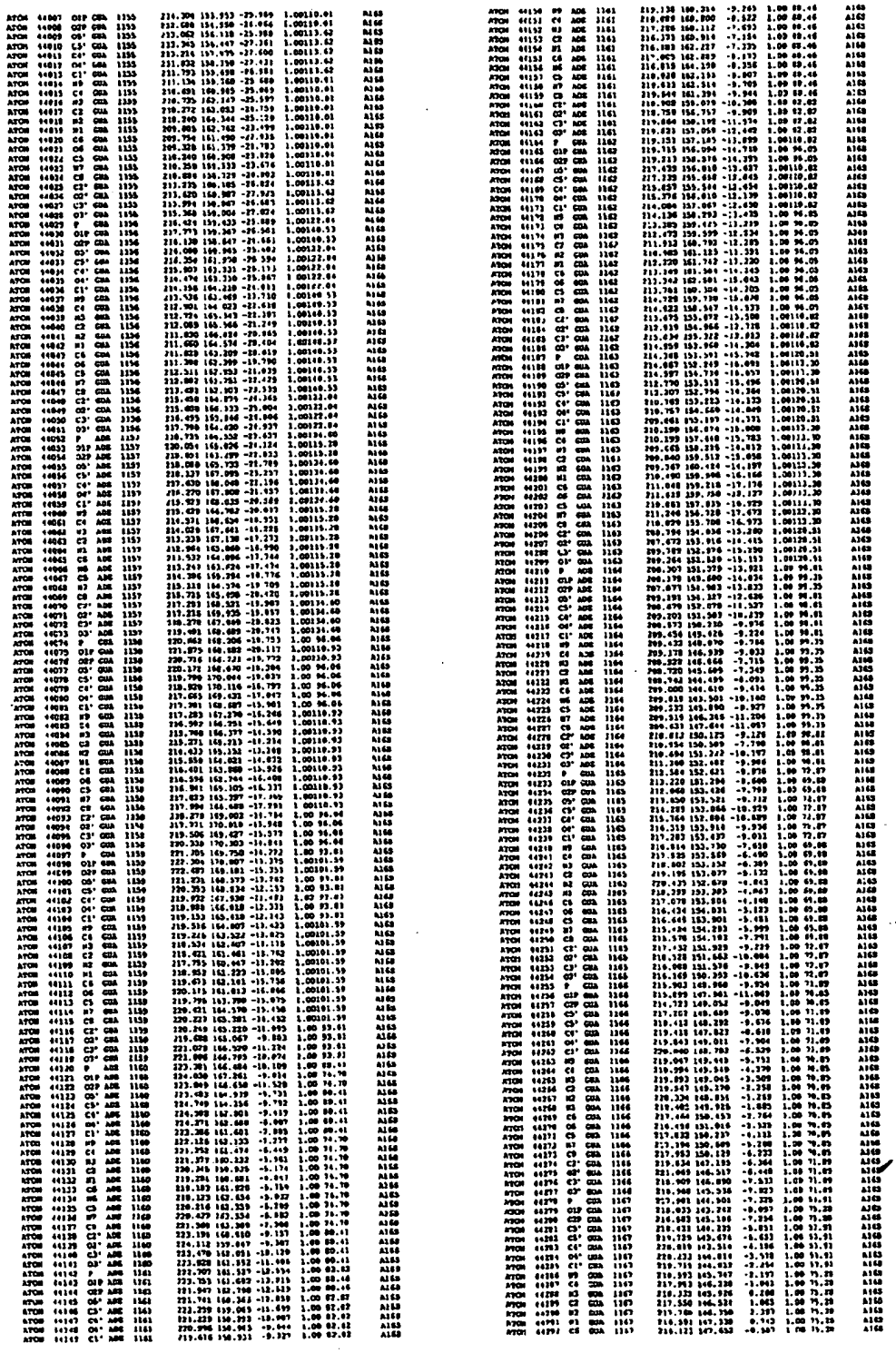




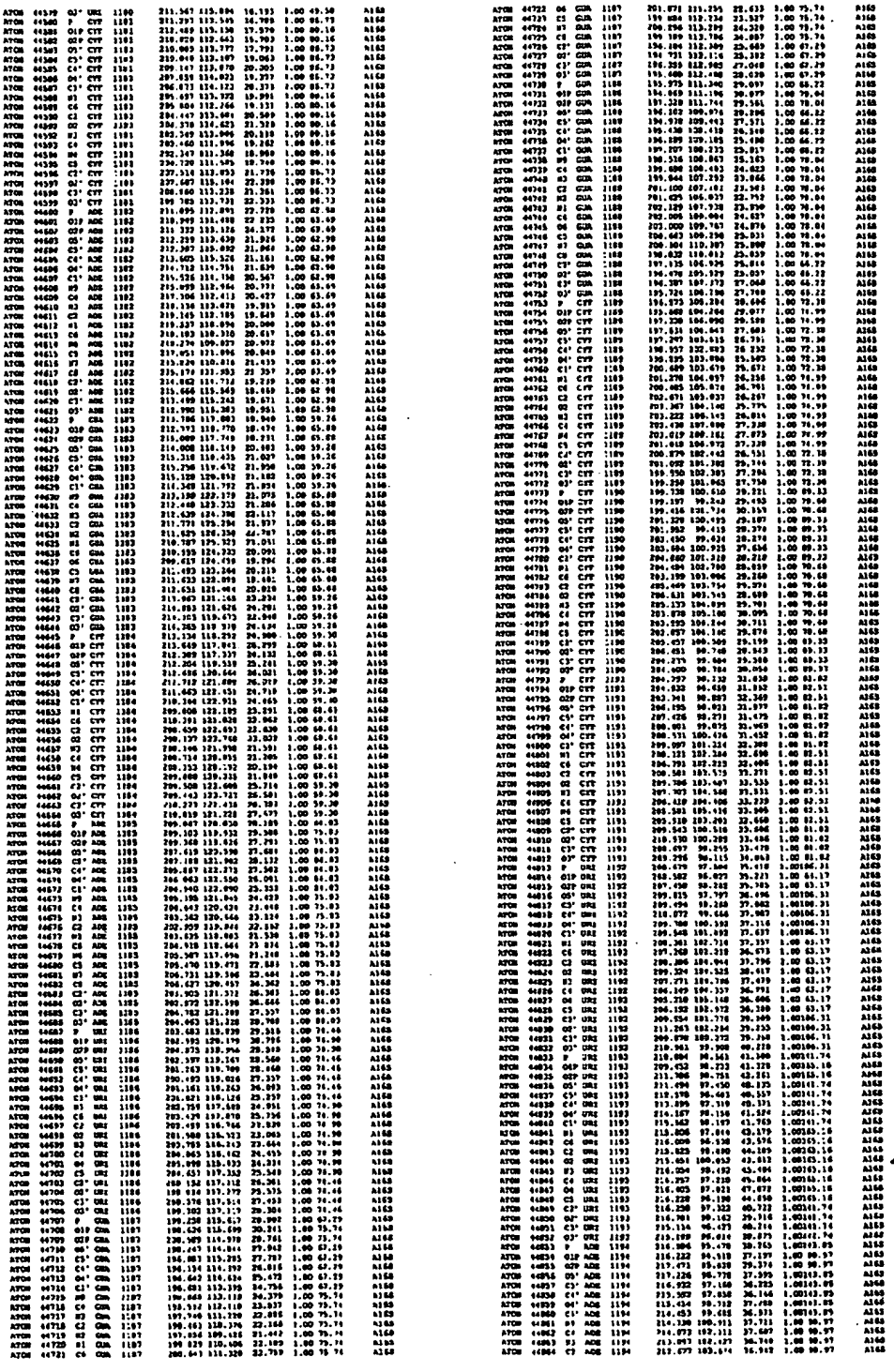


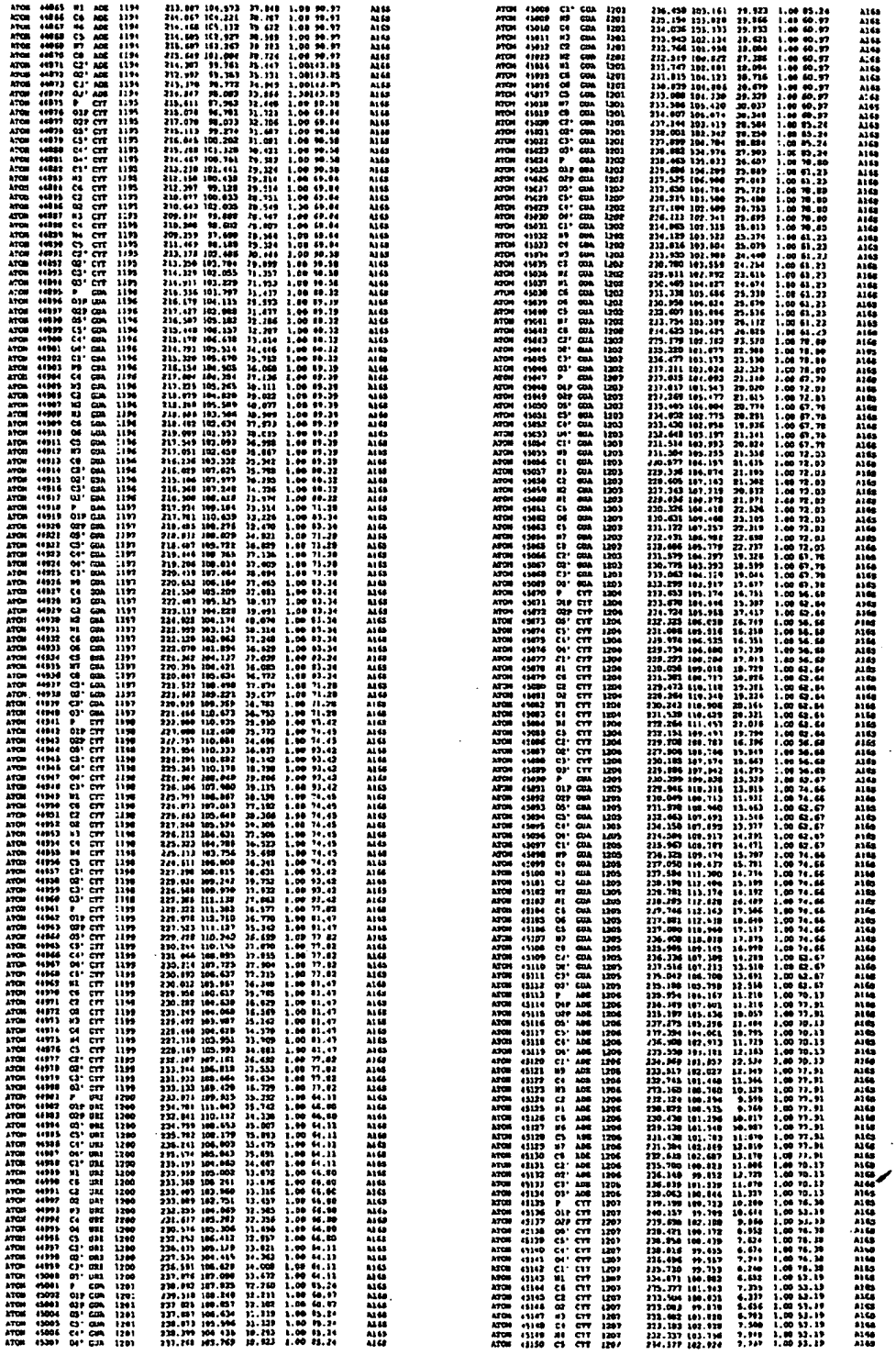


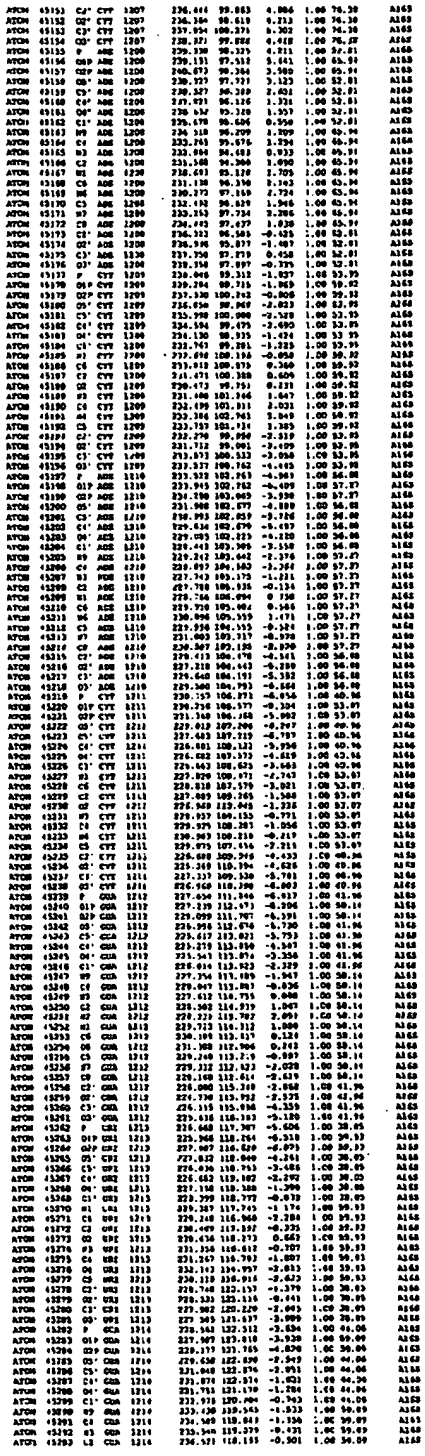


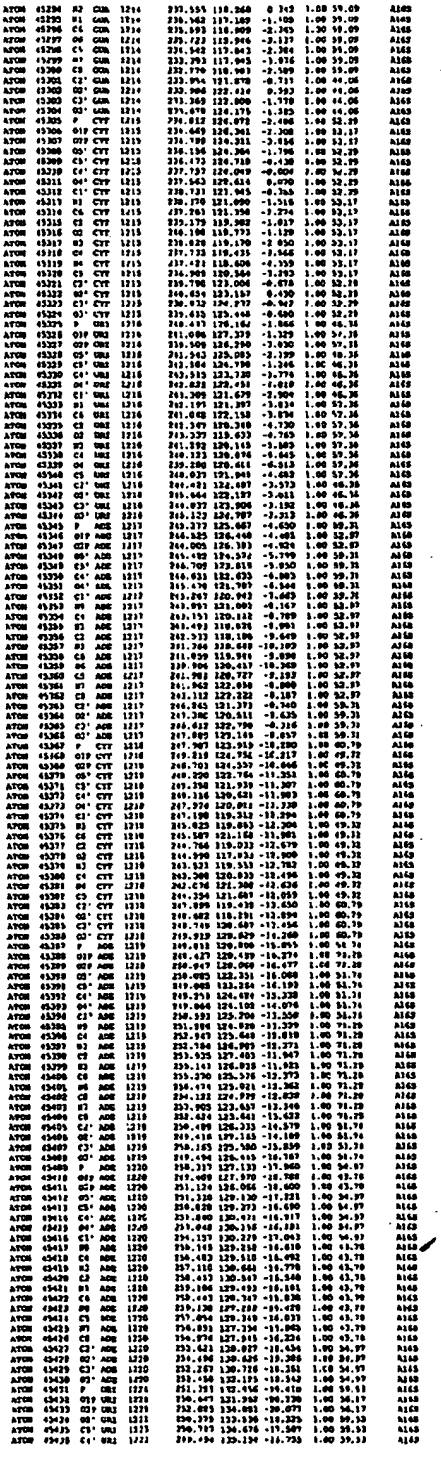




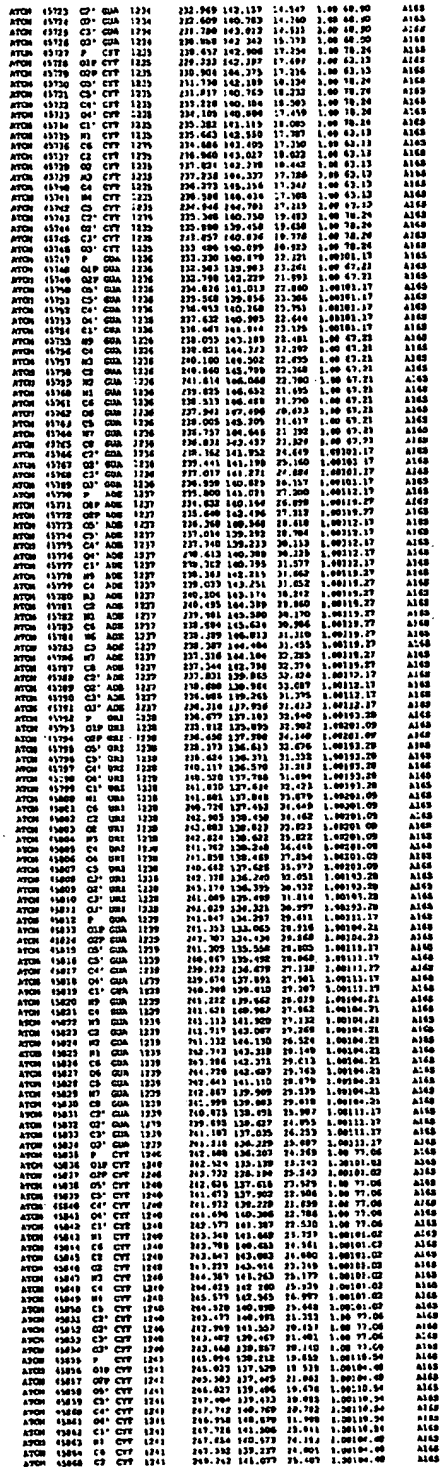


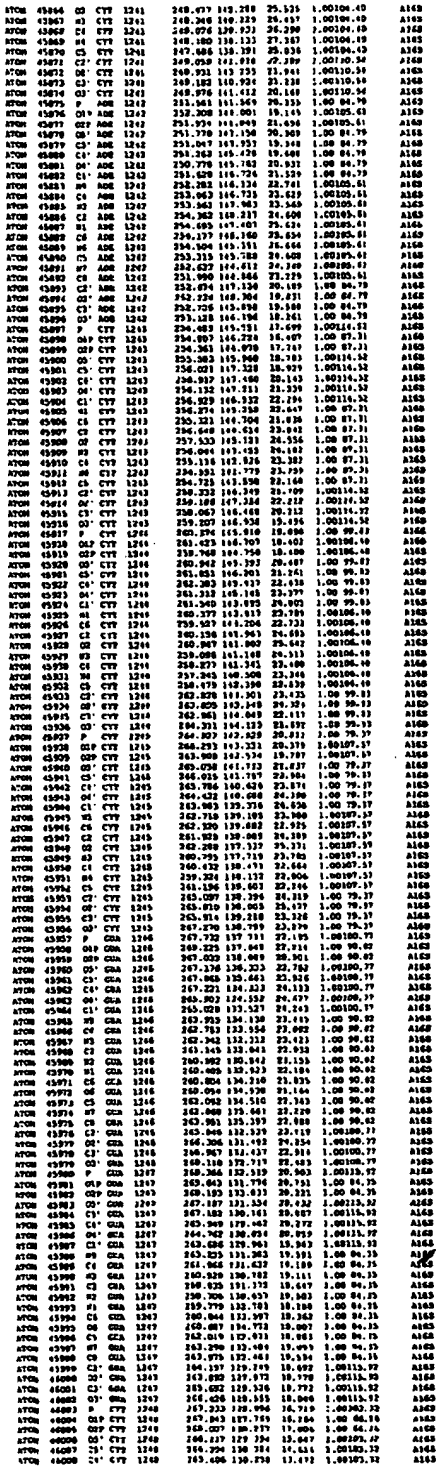


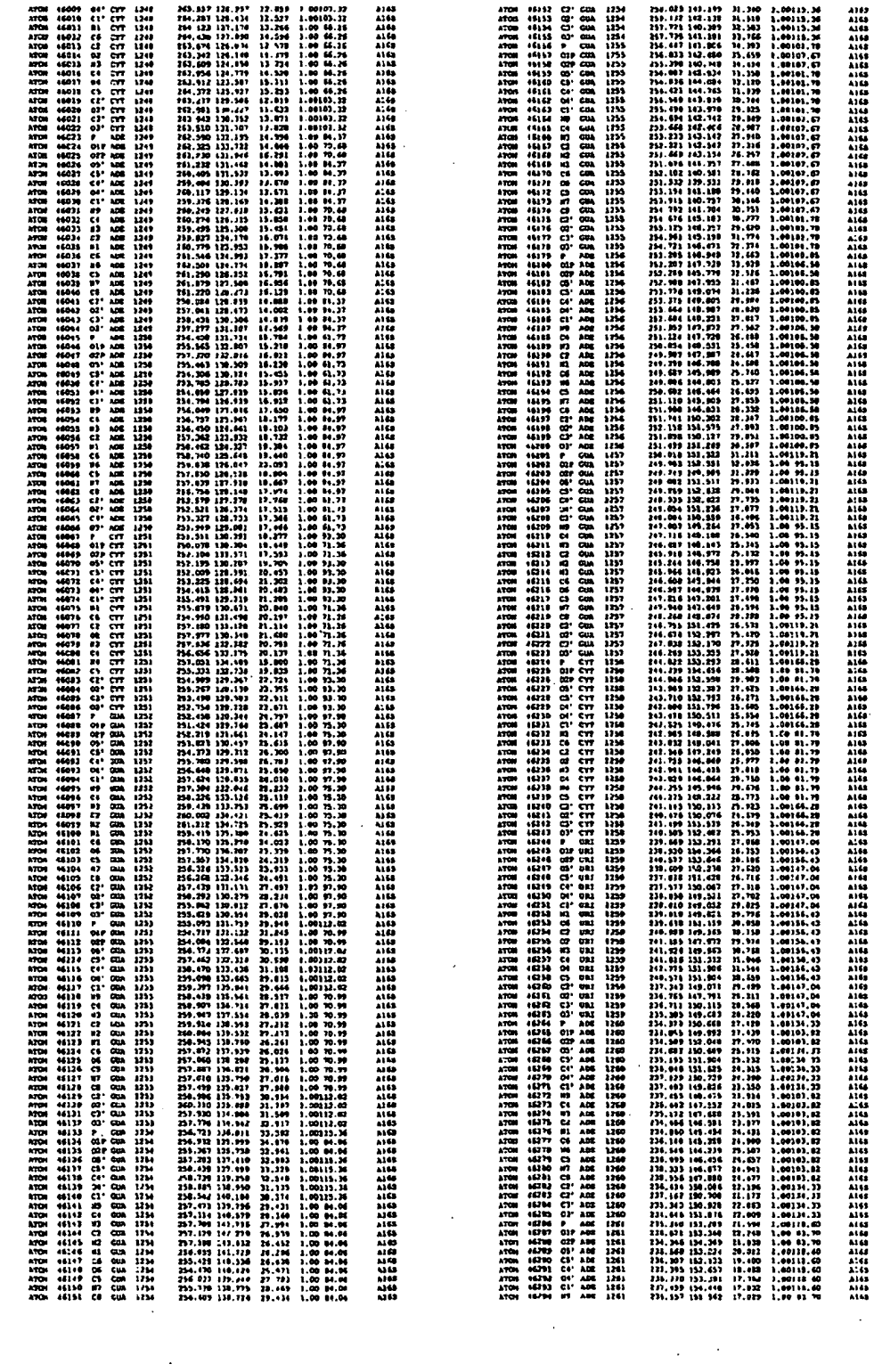


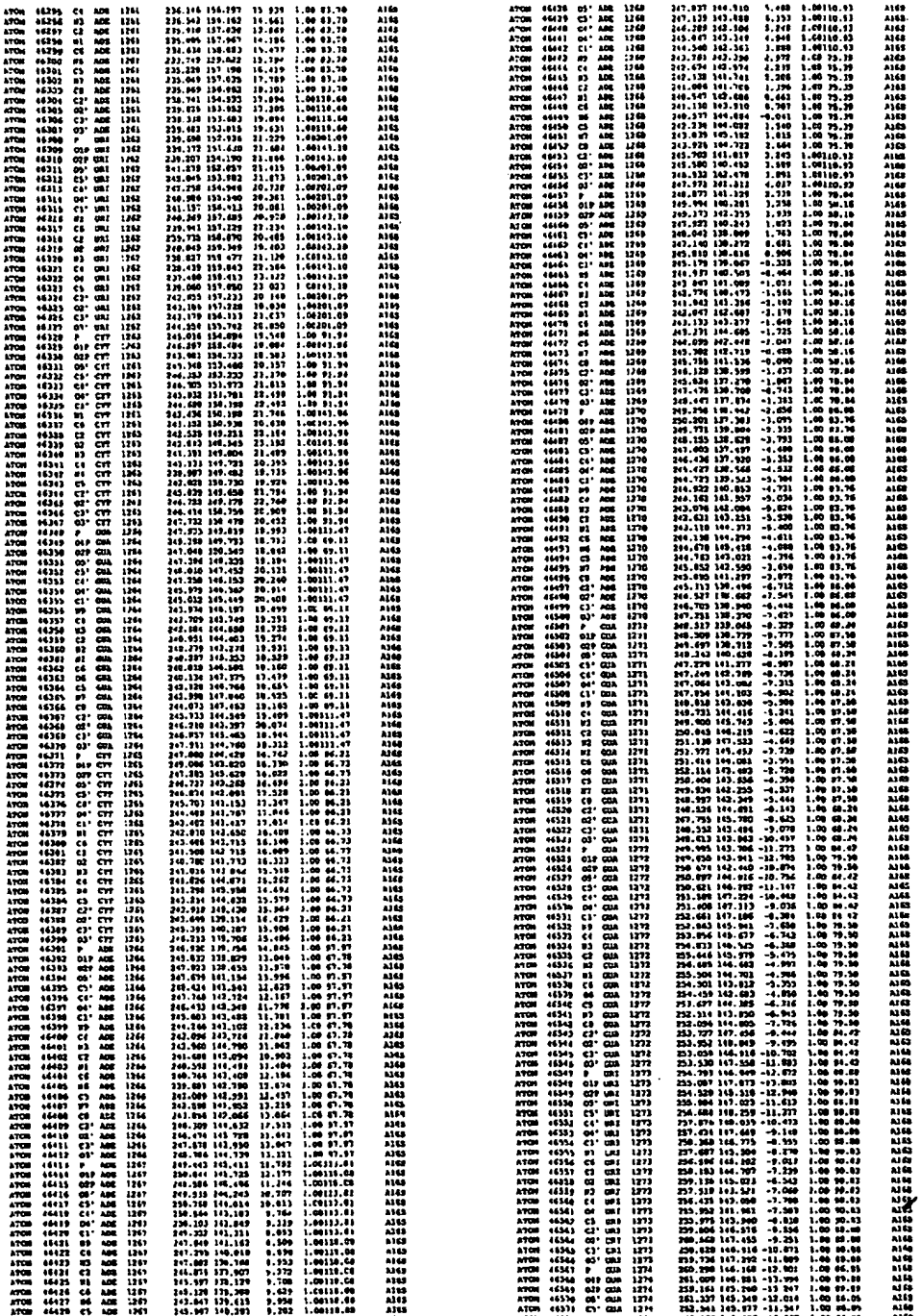


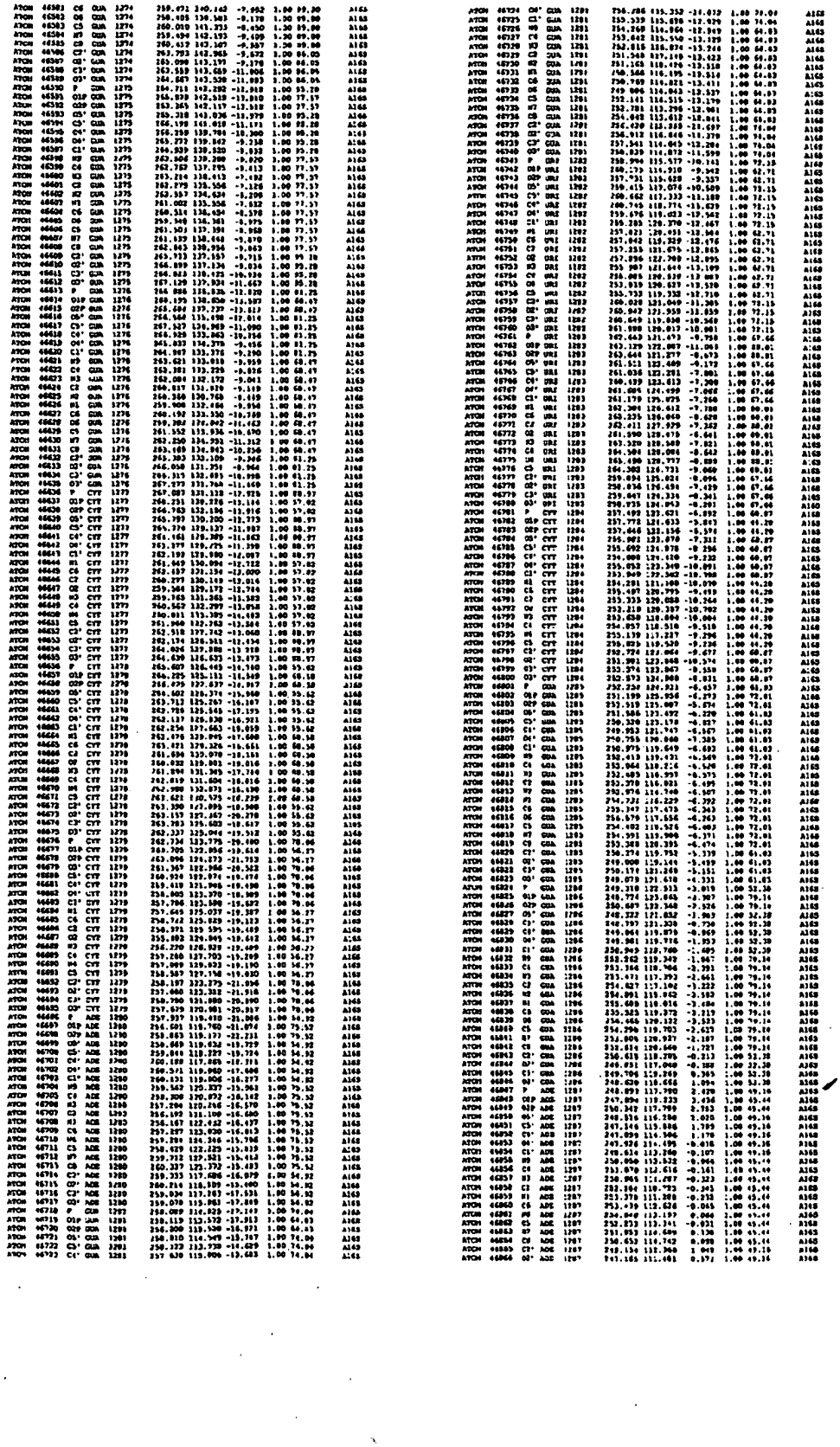


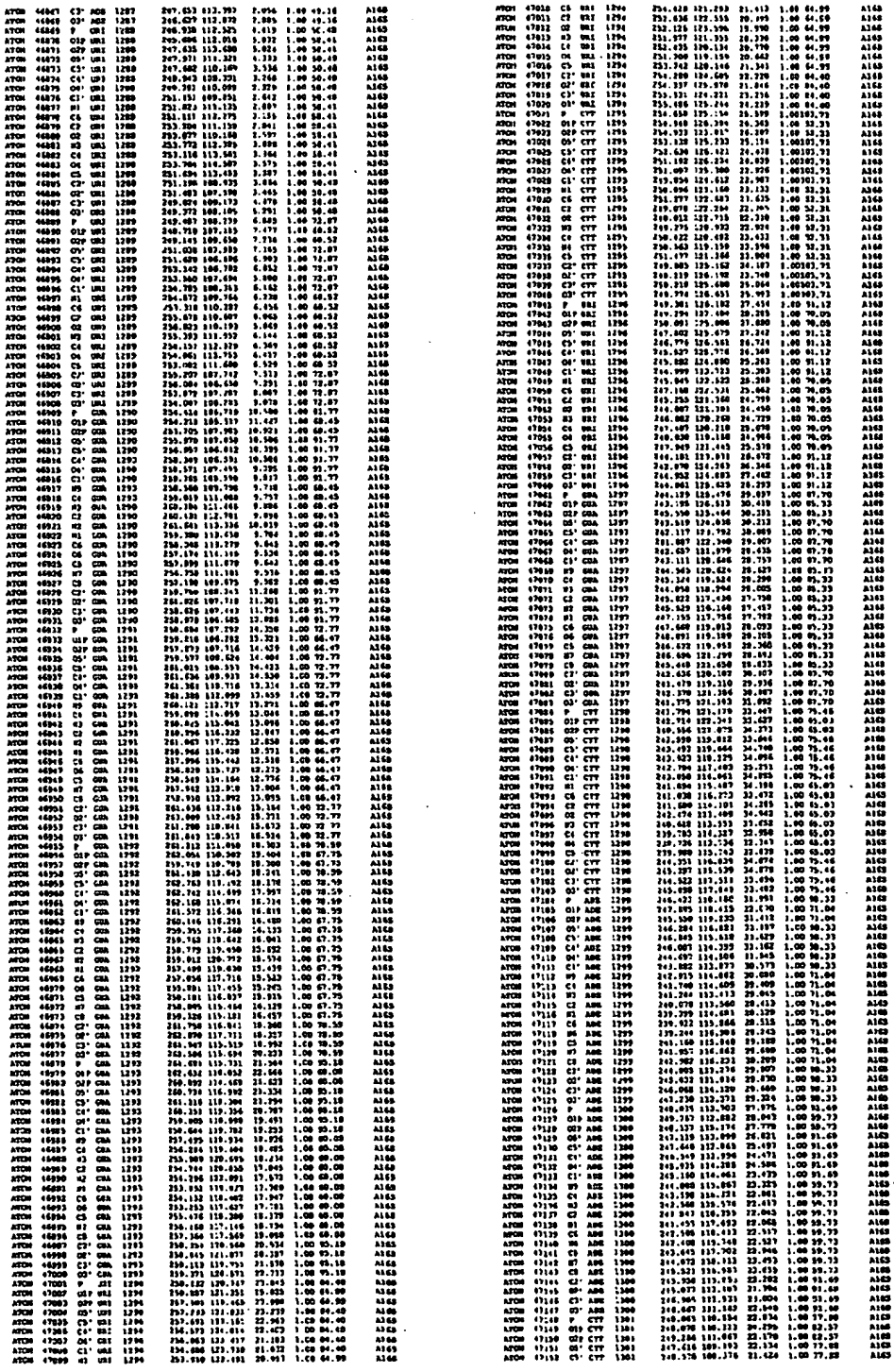




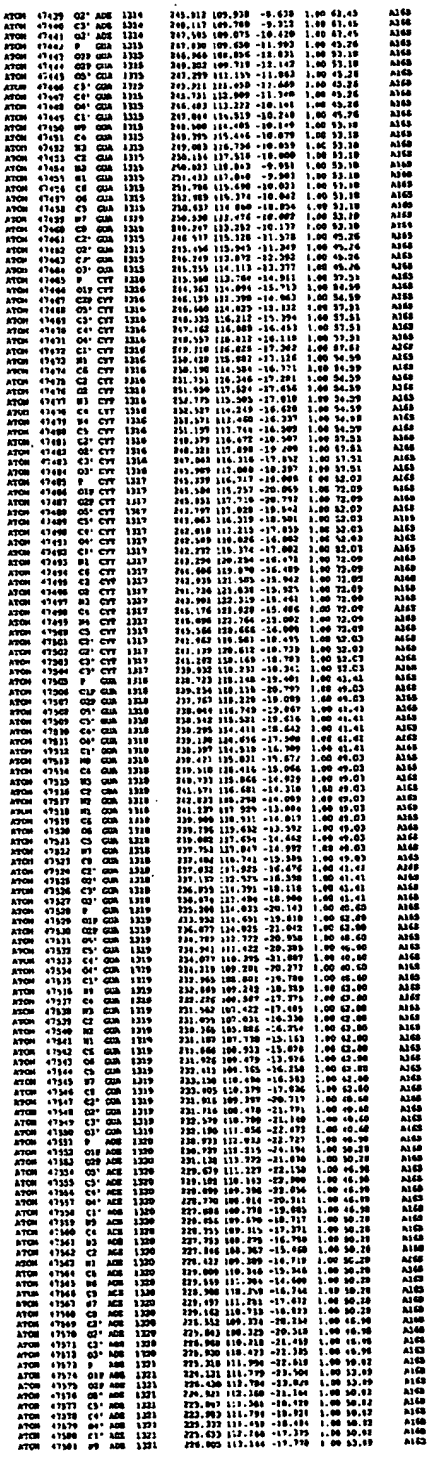


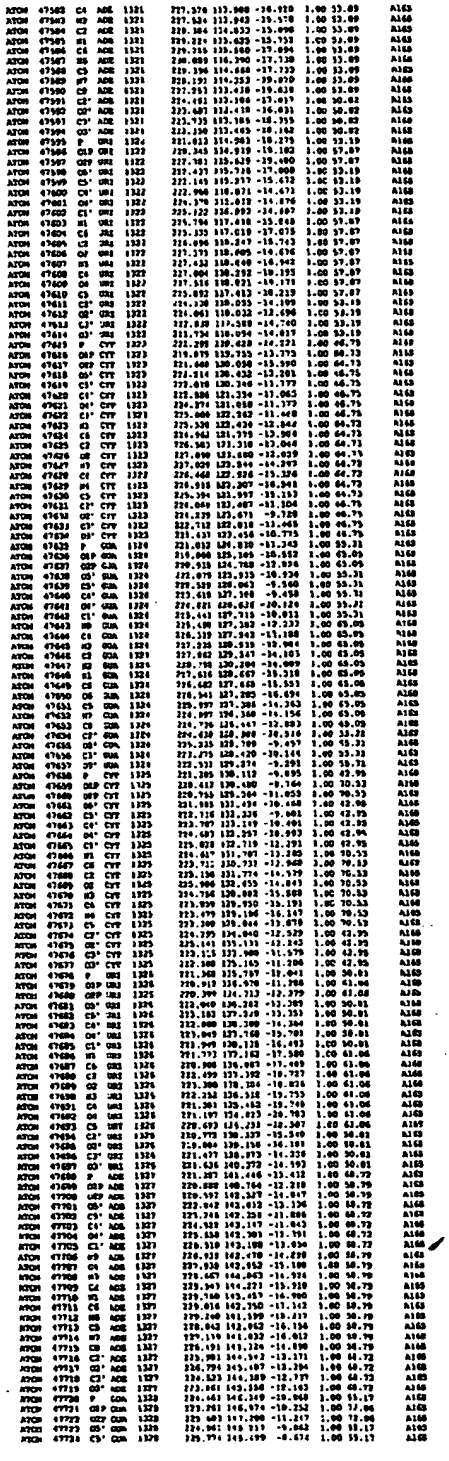


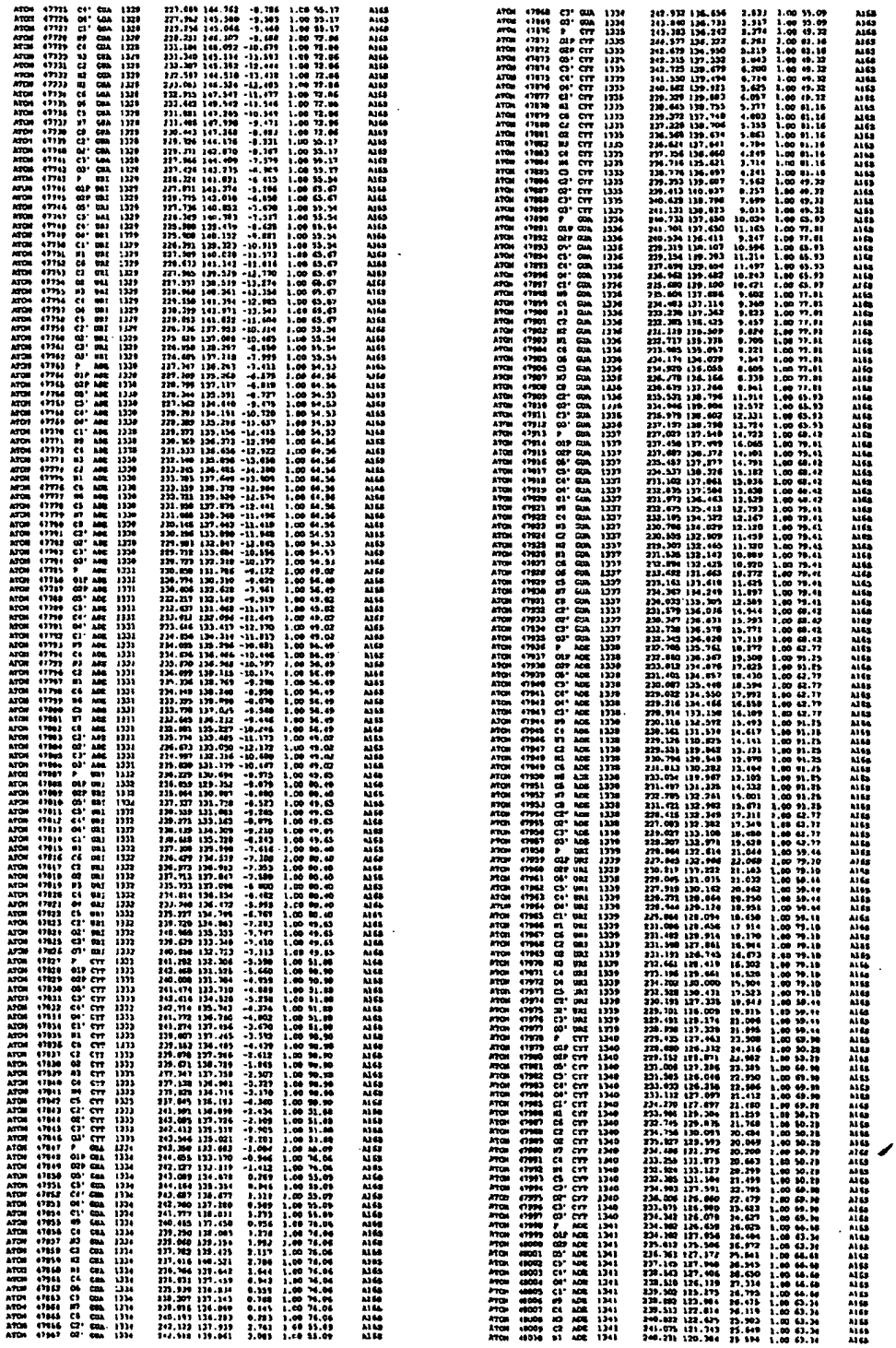


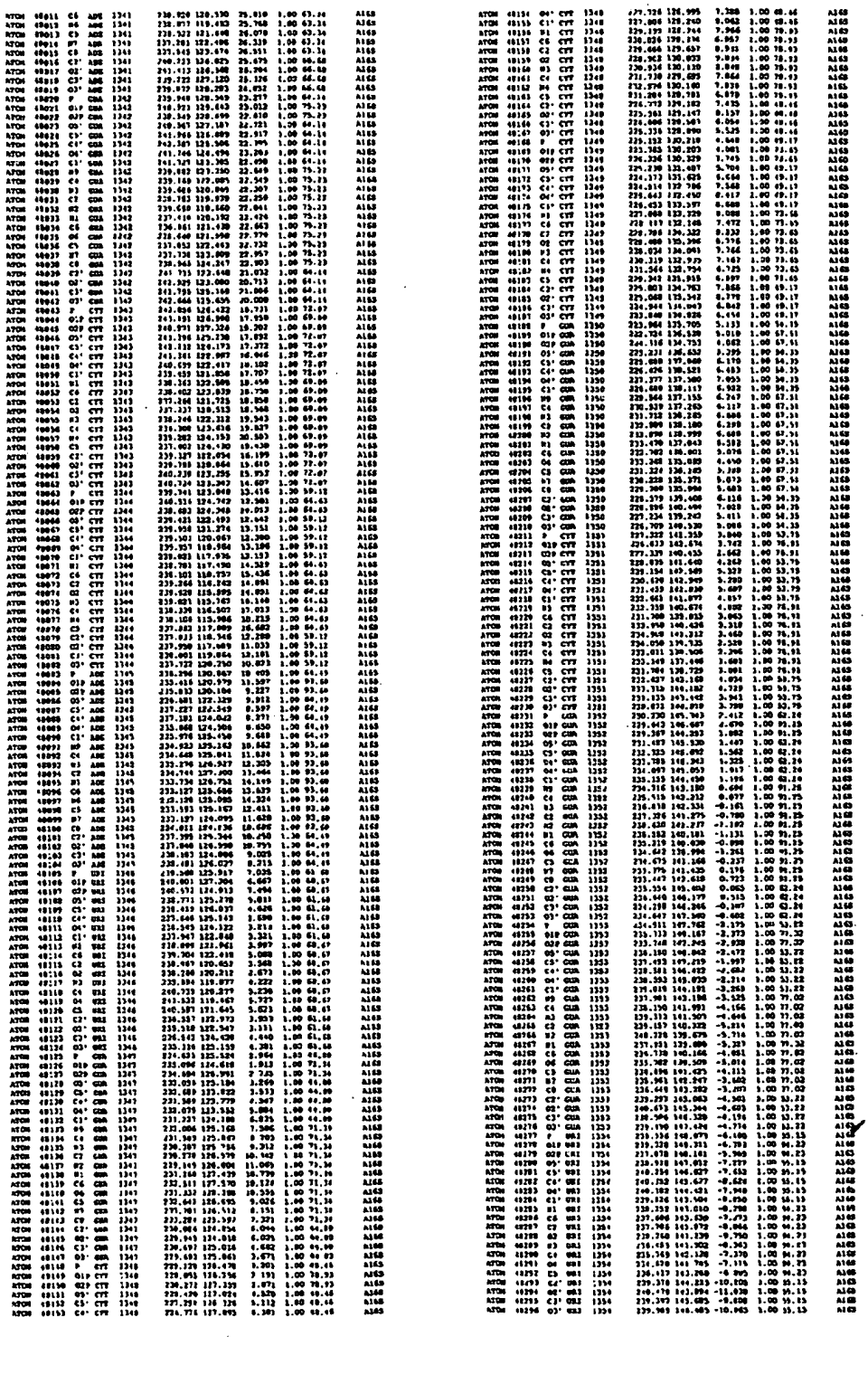


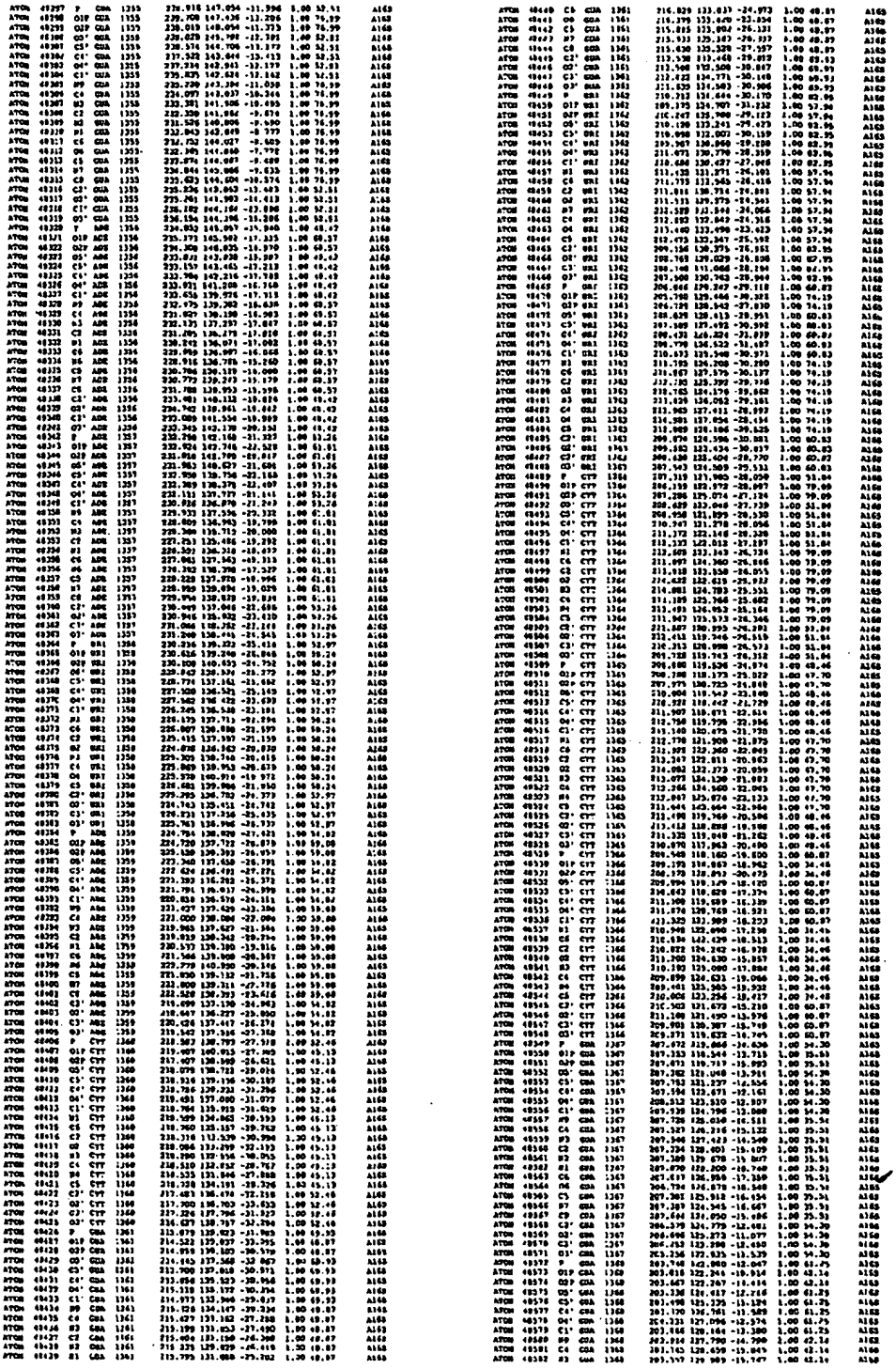


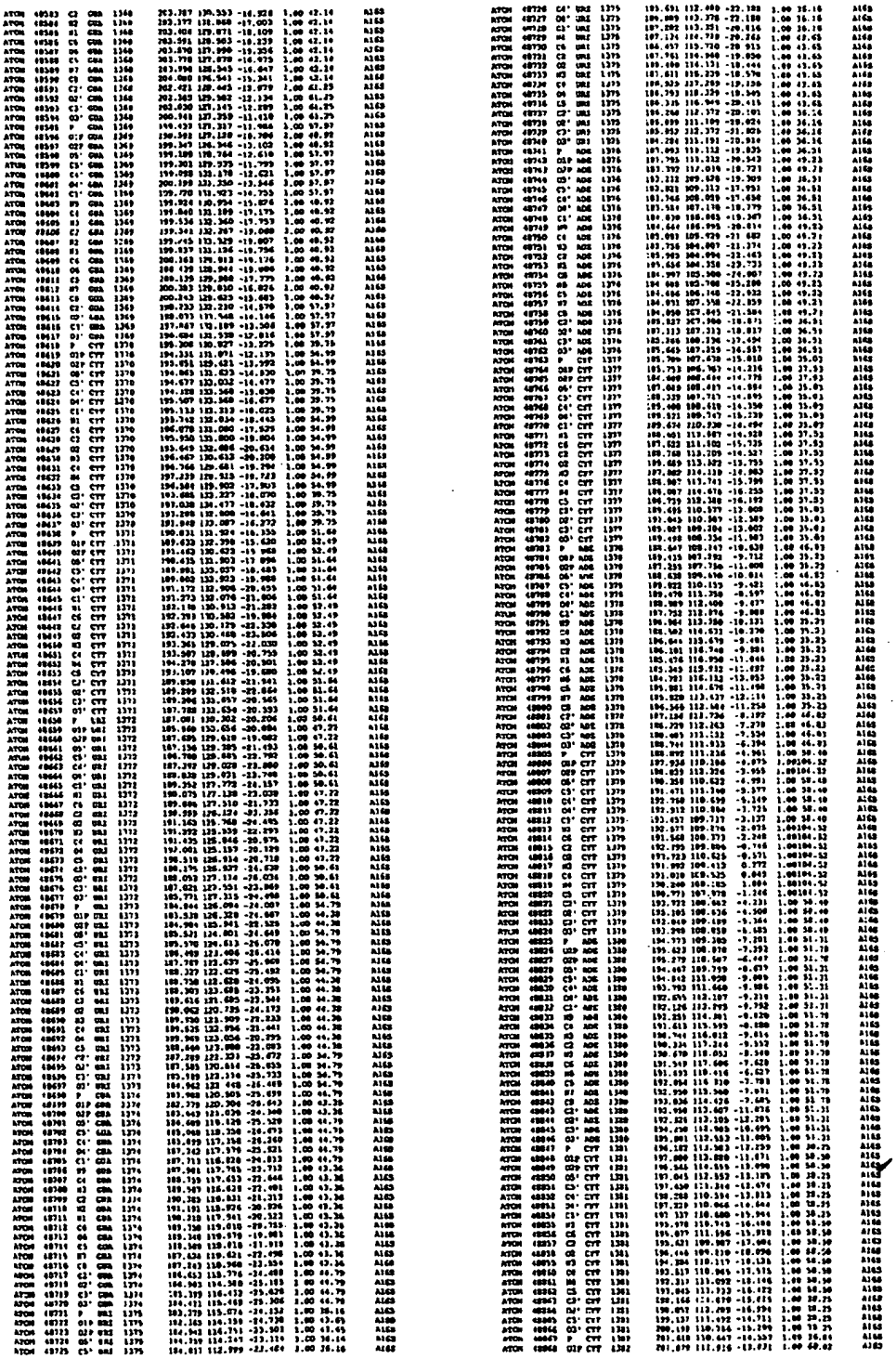


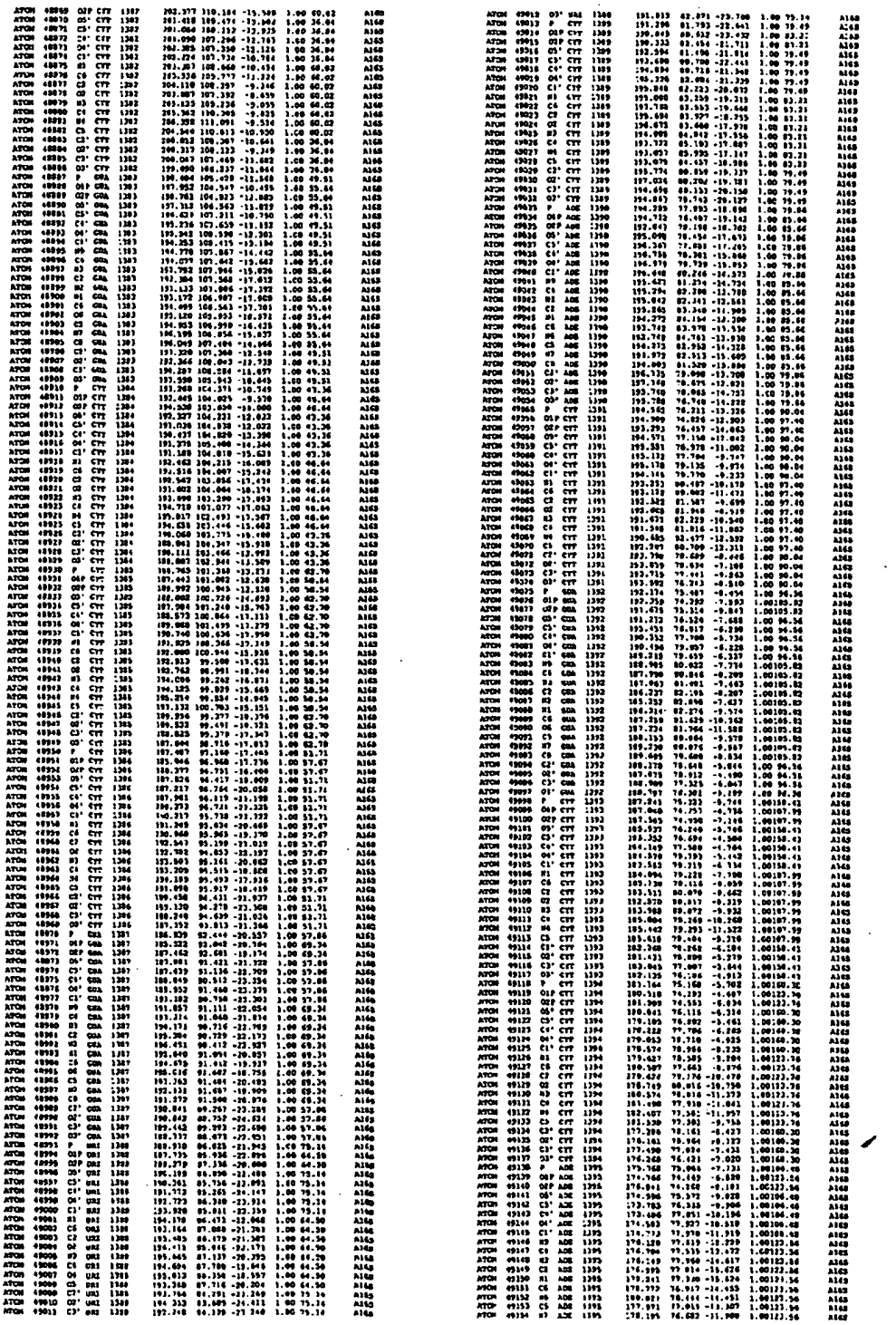


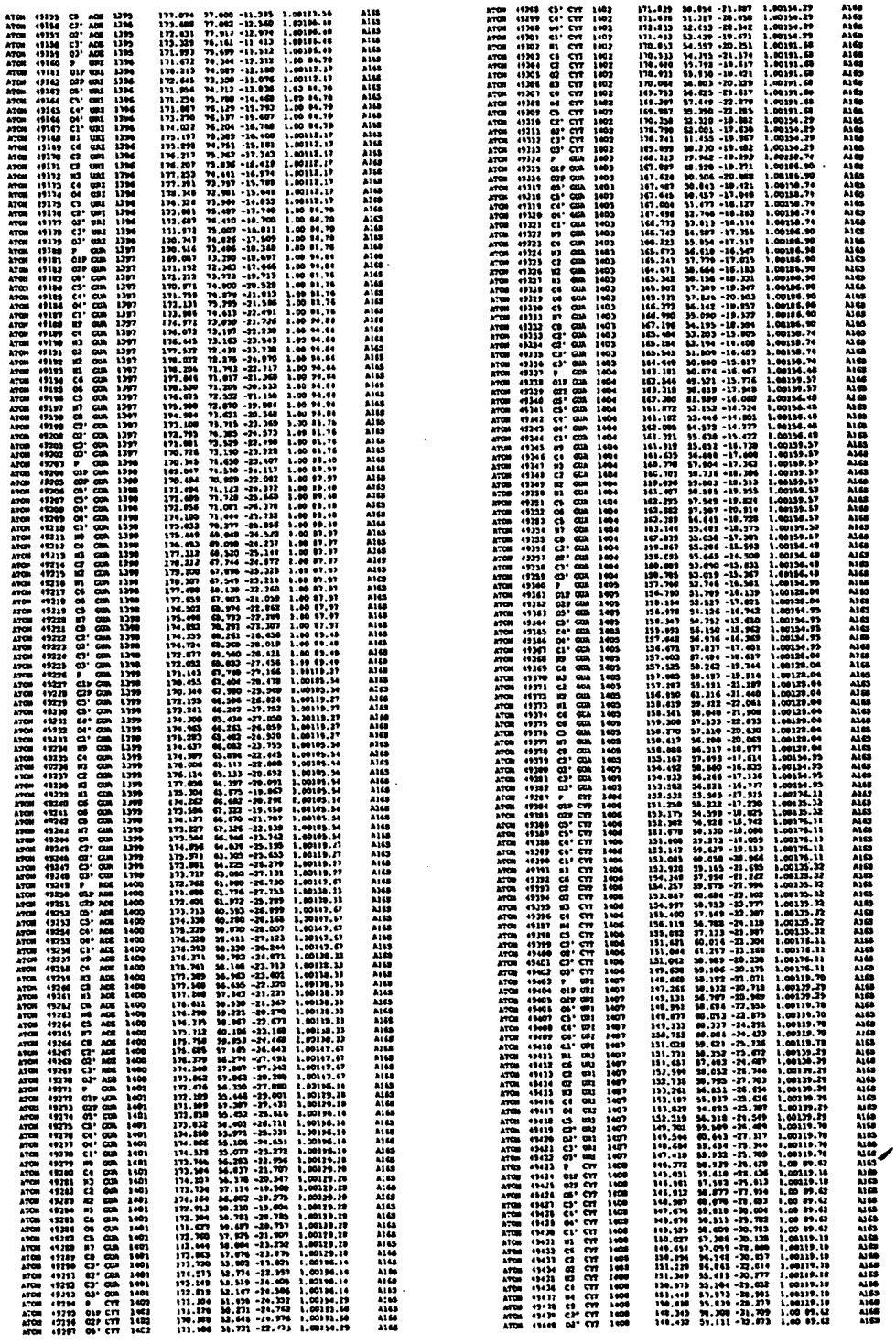


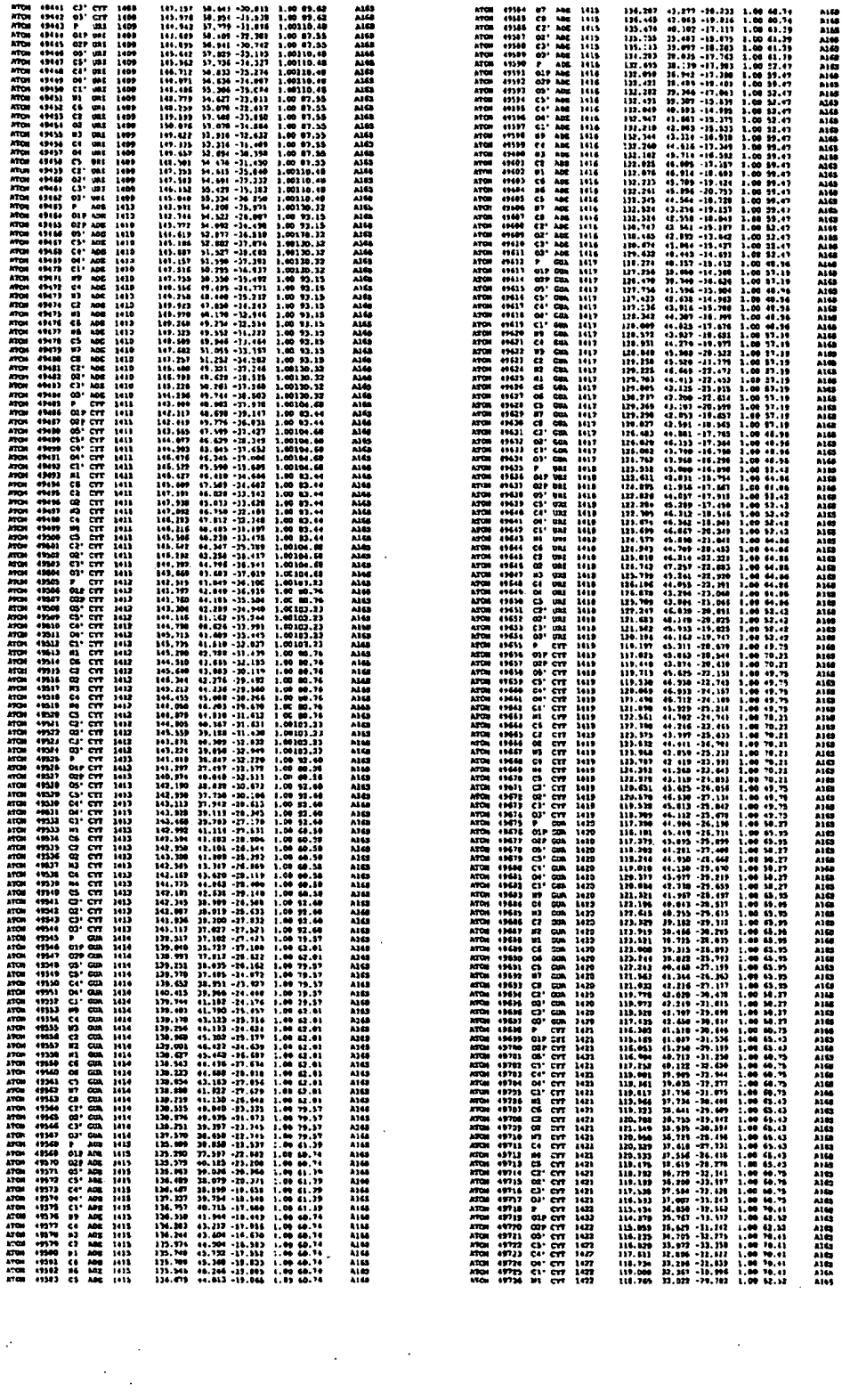




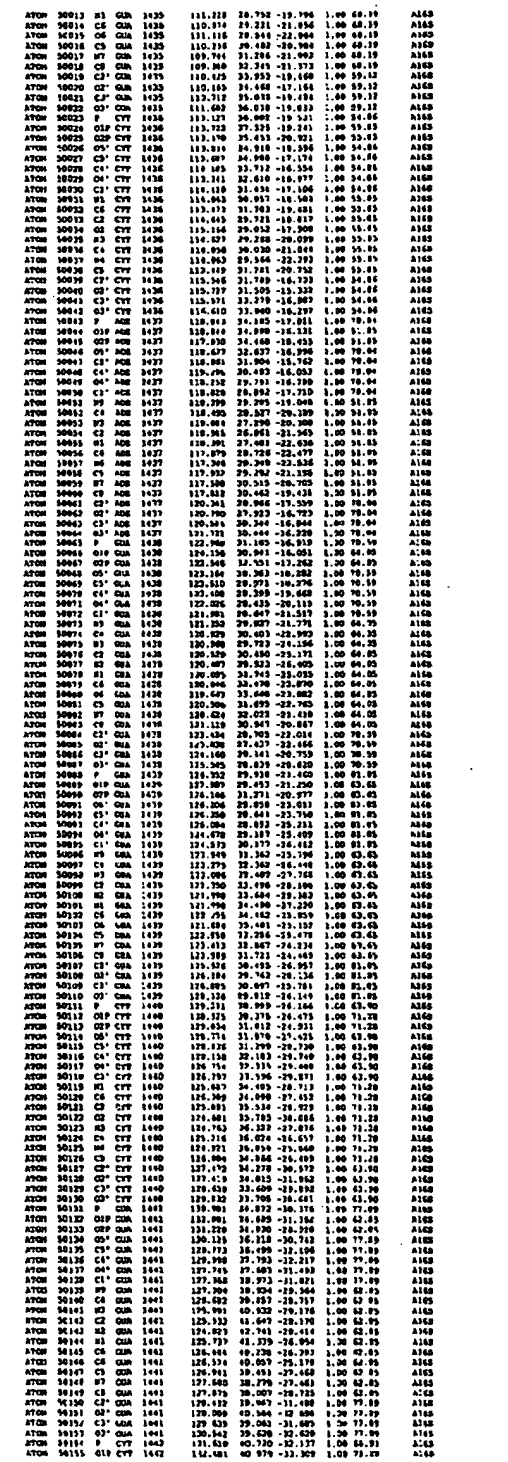


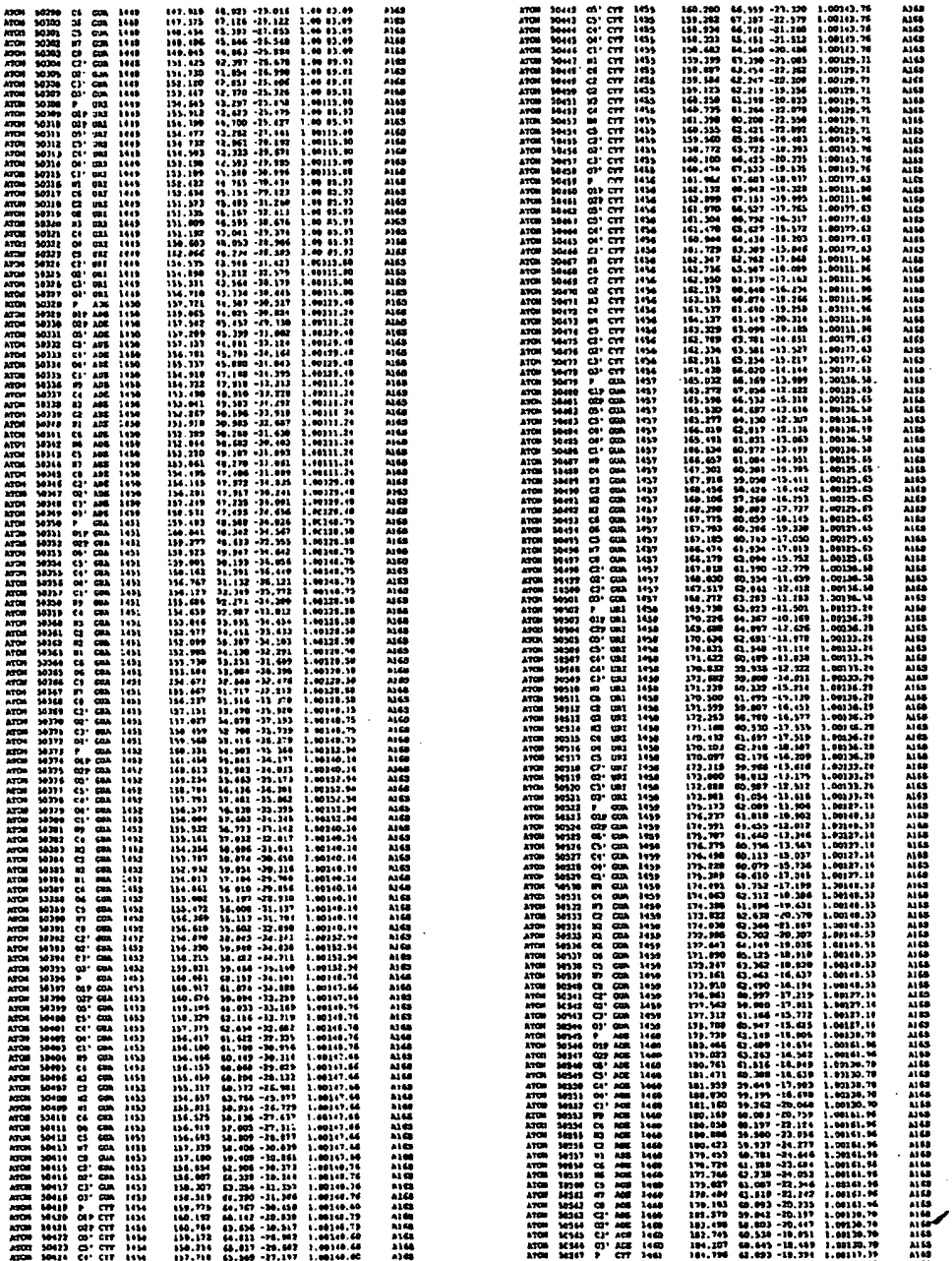


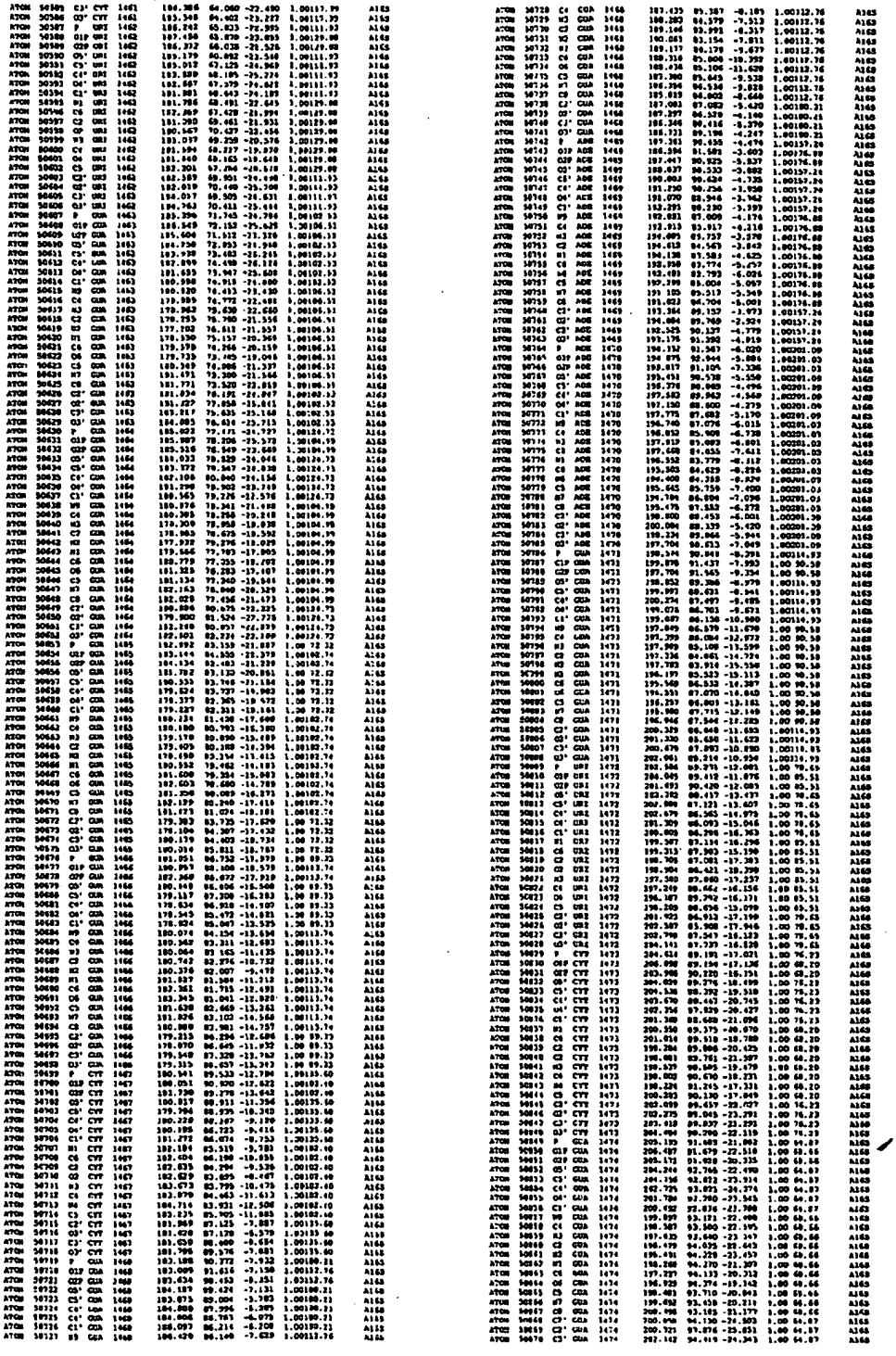


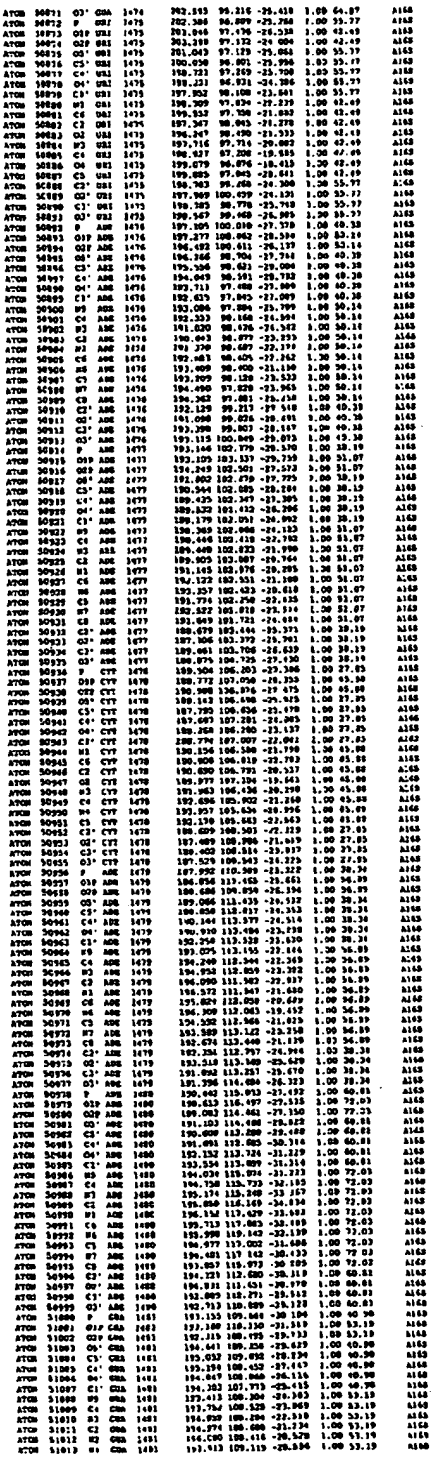


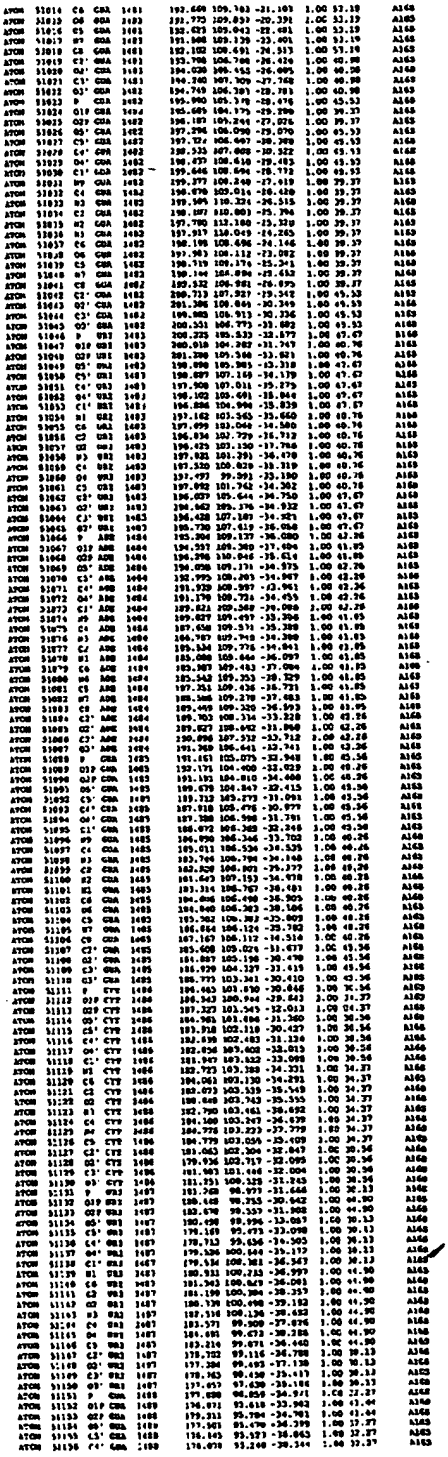


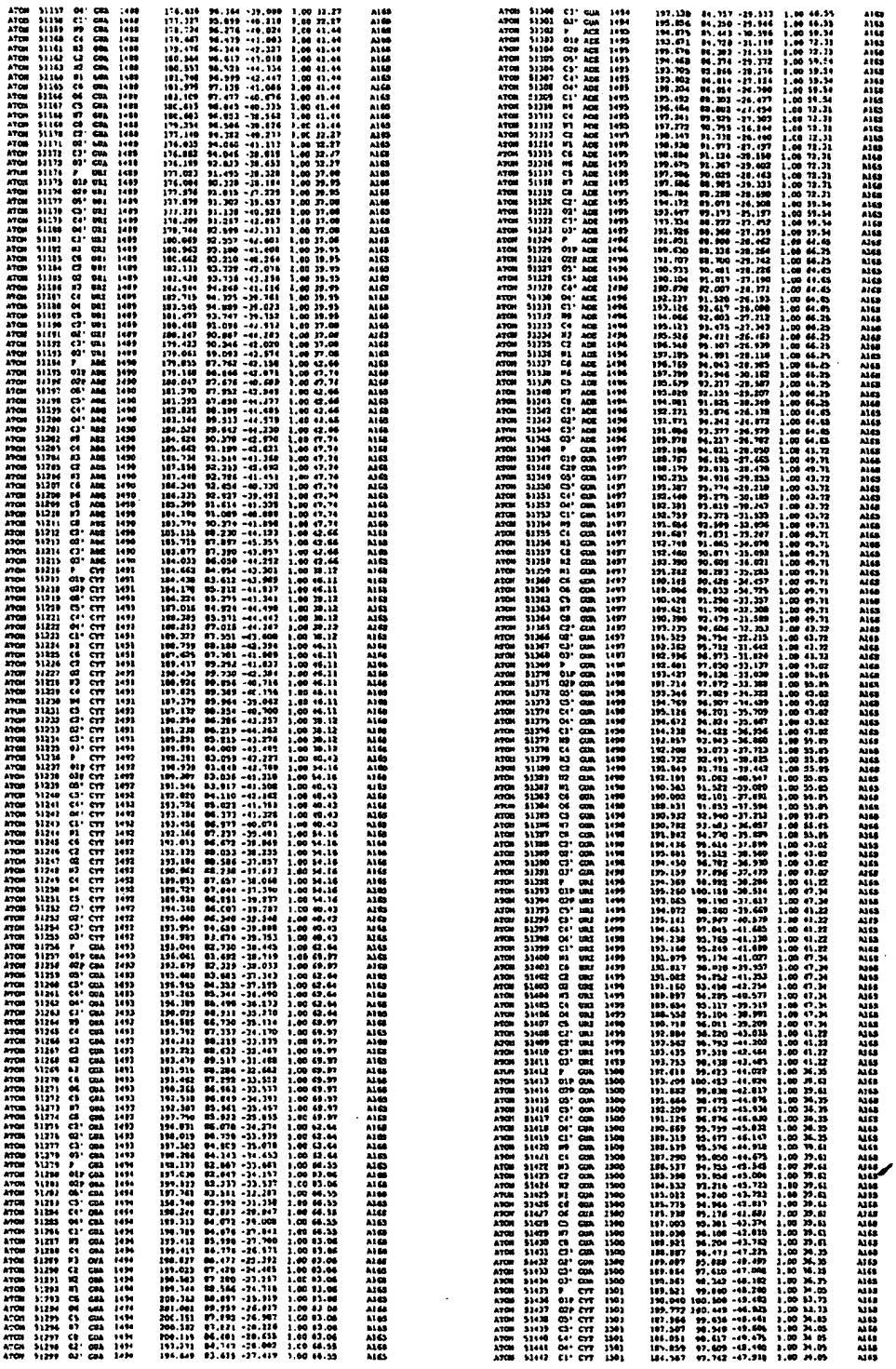


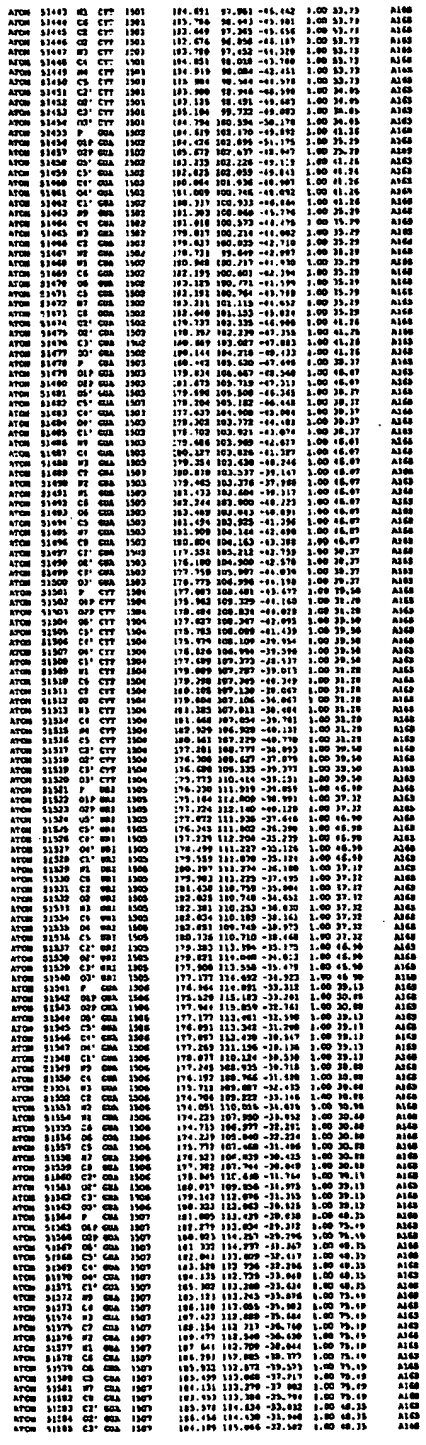


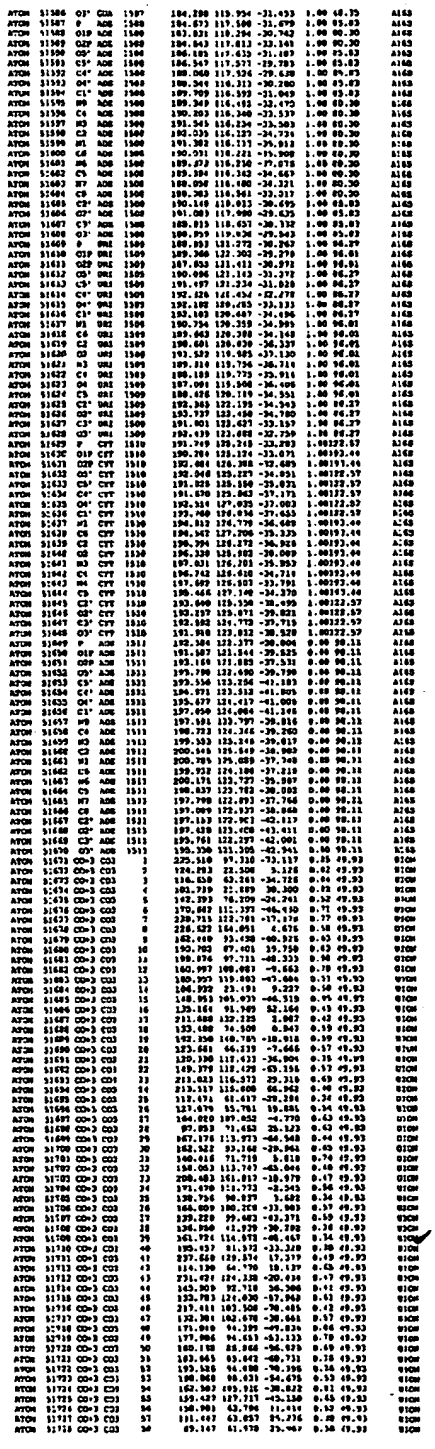












```
ATOM  51728 CD+3 CD3   59     166.821 109.189  17.068  0.71 49.93      3ICM
ATOM  51730 CD+3 CD3   60     159.610  99.717  -8.807  0.68 49.93      3ICM
ATOM  51731 CD+3 CD3   61     179.076 108.850  -2.382  0.61 49.93      3ICM
ATOM  51732 CD+3 CD3   62     173.932 104.934 -25.312  0.57 49.93      3ICM
ATOM  51733 CD+3 CD3   63     159.181 105.993 -16.478  0.78 49.93      3ICM
ATOM  51734 CD+3 CD3   64     164.235  68.008   3.206  0.61 49.93      3ICM
ATOM  51735 CD+3 CD3   65     118.149 102.398 -26.542  0.65 49.93      3ICM
ATOM  51736 CD+3 CD3   66     158.782 135.761 -59.522  0.39 49.93      3ICM
ATOM  51737 CD+3 CD3   67     152.943 110.432 -35.155  0.33 49.93      3ICM
ATOM  51738 CD+3 CD3   68     163.824 127.240 -33.905  0.66 49.93      3ICM
ATOM  51739 CD+3 CD3   69     168.661 112.748 -25.879  0.33 49.93      3ICM
ATOM  51740 CD+3 CD3   70     157.359 116.176 -16.180  0.68 49.93      3ICM
ATOM  51741 CD+3 CD3   71     160.774 131.683 -23.075  0.69 49.93      3ICM
ATOM  51742 CD+3 CD3   72     113.979  99.898 -21.219  0.36 49.93      3ICM
ATOM  51743 CD+3 CD3   73     168.165  98.638 -12.496  0.56 49.93      3ICM
ATOM  51744 CD+3 CD3   74     204.880 123.730 -19.541  0.62 49.93      3ICM
ATOM  51745 CD+3 CD3   75     222.570 176.909 -21.385  0.75 49.93      3ICM
ATOM  51746 CD+3 CD3   76     217.186 130.236 -12.689  0.62 49.93      3ICM
ATOM  51747 CD+3 CD3   77     211.256 109.023  17.689  0.47 49.93      3ICM
ATOM  51748 CD+3 CD3   78     237.742 111.913   0.994  0.31 49.93      3ICM
ATOM  51749 CD+3 CD3   79     232.658 139.365  -2.834  0.63 49.93      3ICM
ATOM  51750 CD+3 CD3   80     121.008  91.484   9.711  0.41 49.93      3ICM
ATOM  51751 CD+3 CD3   81     166.656 131.932  -7.282  0.65 49.93      3ICM
ATOM  51752 CD+3 CD3   82     127.818  96.504  26.317  0.62 49.93      3ICM
ATOM  51753 CD+3 CD3   83     116.653  97.672 -12.963  0.41 49.93      3ICM
ATOM  51754 CD+3 CD3   84     114.665  92.980  -3.199  0.26 49.93      3ICM
ATOM  51755 CD+3 CD3   85     208.971 102.629 -75.378  0.47 49.93      3ICM
ATOM  51756 CD+3 CD3   86     138.375 111.795 -37.162  0.57 49.93      3ICM
ATOM  51757 CD+3 CD3   87     120.597 118.767 -54.803  0.46 49.93      3ICM
ATOM  51758 CD+3 CD3   88     235.783 113.754  -7.199  0.71 49.93      3ICM
ATOM  51759 CD+3 CD3   89     237.982 176.033  -6.389  0.35 49.93      3ICM
ATOM  51760 CD+3 CD3   90     209.515 110.511  10.517  0.71 49.93      3ICM
ATOM  51761 CD+3 CD3   91     236.983 119.764  25.349  0.51 49.93      3ICM
ATOM  51762 CD+3 CD3   92     233.335 121.967  26.400  0.68 49.93      3ICM
ATOM  51763 CD+3 CD3   93     214.391 126.842  82.280  0.23 49.93      3ICM
ATOM  51764 CD+3 CD3   94     243.857 105.664  47.628  0.36 49.93      3ICM
ATOM  51765 CD+3 CD3   95     202.116 111.973  10.190  0.32 49.93      3ICM
ATOM  51766 CD+3 CD3   96     201.716 111.648   7.198  0.63 49.93      3ICM
ATOM  51767 CD+3 CD3   97     252.927 113.304   7.891  0.38 49.93      3ICM
ATOM  51768 CD+3 CD3   98     260.686 139.389 -11.185  0.26 49.93      3ICM
ATOM  51769 CD+3 CD3   99     163.090  93.335 -33.946  0.31 49.93      3ICM
ATOM  51770 CD+3 CD3  100     112.583  58.260 -26.742  0.47 49.93      3ICM
ATOM  51771 CD+3 CD3  101     120.953  35.446 -20.904  0.33 49.93      3ICM
ATOM  51772 CD+3 CD3  102     128.309  40.125 -21.897  0.35 49.93      3ICM
ATOM  51773 CD+3 CD3  103     141.210  49.206 -34.818  0.27 49.93      3ICM
ATOM  51774 CD+3 CD3  104     171.587  67.035 -34.550  0.48 49.93      3ICM
ATOM  51775 CD+3 CD3  107     177.941 108.398 -12.763  0.68 49.93      3ICM
ATOM  51776 CD+3 CD3  108      98.701  25.921  -0.795  0.18 49.93      3ICM
ATOM  51777 CD+3 CD3  109      86.428  44.644 -21.125  0.29 49.93      3ICM
ATOM  51778 CD+3 CD3  110     139.990  57.474 -41.900  0.18 49.93      3ICM
ATOM  51779 CD+3 CD3  111     138.183  69.411 -43.171  0.29 49.93      3ICM
ATOM  51780 CD+3 CD3  112     138.606  95.111  -7.895  0.35 49.93      3ICM
ATOM  51781 CD+3 CD3  113     147.694 102.895  -3.972  0.54 49.93      3ICM
ATOM  51782 CD+3 CD3  114     114.218  51.772 -12.537  0.64 49.93      3ICM
ATOM  51783 CD+3 CD3  115     139.573  69.638   9.967  0.58 49.93      3ICM
ATOM  51784 CD+3 CD3  116     153.253  93.282  46.388  0.29 49.93      3ICM
ATOM  51785 CD+3 CD3  117     186.908 101.314  22.601  0.54 49.93      3ICM
ATOM  51786 CD+3 CD3  118     168.915 104.813 -25.911  0.35 49.93      3ICM
ATOM  51787 CD+3 CD3  119     135.698 110.566 -46.923  0.56 49.93      3ICM
ATOM  51788 CD+3 CD3  120     112.151 109.944 -32.974  0.45 49.93      3ICM
ATOM  51789 CD+3 CD3  121     122.564  90.091   2.953  0.47 49.93      3ICM
ATOM  51790 CD+3 CD3  122     218.798  93.898 -50.980  0.55 49.93      3ICM
ATOM  51791 CD+3 CD3  123      97.173  38.538  13.525  0.26 49.93      3ICM
ATOM  51792 CD+3 CD3  124     168.186  92.879 -20.348  0.29 49.93      3ICM
ATOM  51793 CD+3 CD3  125     168.049  78.751 -33.260  0.44 49.93      3ICM
ATOM  51794 CD+3 CD3  126     167.213  74.605 -43.389  0.32 49.93      3ICM
ATOM  51795 CD+3 CD3  127     143.687  58.842   0.976  0.43 49.93      3ICM
ATOM  51796 CD+3 CD3  128     211.296 118.978  -1.556  0.26 49.93      3ICM
ATOM  51797 CD+3 CD3  129     205.227 138.919  -7.894  0.34 49.93      3ICM
ATOM  51798 CD+3 CD3  130     204.581 143.900   8.062  0.57 49.93      3ICM
ATOM  51799 CD+3 CD3  131     169.253 126.650 -18.195  0.48 49.93      3ICM
ATOM  51800 CD+3 CD3  132     178.992 125.033 -16.333  0.68 49.93      3ICM
ATOM  51801 CD+3 CD3  133     197.382 136.884 -14.695  0.68 49.93      3ICM
ATOM  51802 CD+3 CD3  134     194.668 146.886  -5.548  0.51 49.93      3ICM
ATOM  51803 CD+3 CD3  135     183.984 133.432 -18.584  0.38 49.93      3ICM
ATOM  51804 CD+3 CD3  136     141.802  73.398 -21.340  0.43 49.93      3ICM
ATOM  51805 CD+3 CD3  137     228.752 113.781  13.447  0.63 49.93      3ICM
ATOM  51806 CD+3 CD3  138     263.987 135.498  17.651  0.37 49.93      3ICM
ATOM  51807 CD+3 CD3  139     125.958  69.039 -17.434  0.32 49.93      3ICM
ATOM  51808 CD+3 CD3  140      83.829  47.873   9.791  0.29 49.93      3ICM
ATOM  51809 CD+3 CD3  141     169.289  74.451 -15.908  0.34 49.93      3ICM
ATOM  51810 CD+3 CD3  142     173.414  89.399 -34.988  0.52 49.93      3ICM
ATOM  51811 CD+3 CD3  143     223.489 198.776   8.343  0.41 49.93      3ICM
ATOM  51812 CD+3 CD3  144     222.785 188.023  20.857  0.48 49.93      3ICM
ATOM  51813 CD+3 CD3  145     187.884 184.374  13.727  0.27 49.93      3ICM
ATOM  51814 CD+3 CD3  146     193.482 127.613 -15.821  0.55 49.93      3ICM
ATOM  51815 CD+3 CD3  147     203.425 105.707 -30.484  0.51 49.93      3ICM
ATOM  51816 CD+3 CD3  148     187.885 125.287 -17.581  0.36 49.93      3ICM
ATOM  51817 CD+3 CD3  149     185.374 103.347 -42.065  0.63 49.93      3ICM
ATOM  51818 CD+3 CD3  150     188.486 198.241  30.751  0.23 49.93      3ICM
ATOM  51819 CD+3 CD3  151     161.186  89.373   4.168  0.29 49.93      3ICM
ATOM  51820 CD+3 CD3  152     196.074 103.013 -26.826  0.37 49.93      3ICM
ATOM  51821 CD+3 CD3  153     112.670  61.786  13.572  0.53 49.93      3ICM
ATOM  51822 CD+3 CD3  154     139.087  77.256  23.771  0.56 49.93      3ICM
ATOM  51823 CD+3 CD3  155     157.997  74.310  26.748  0.49 49.93      3ICM
ATOM  51824 CD+3 CD3  156     148.803 104.217   6.779  0.32 49.93      3ICM
ATOM  51825 CD+3 CD3  157     171.311  98.134  22.180  0.38 49.93      3ICM
ATOM  51826 CD+3 CD3  158      96.897  63.765  -1.802  0.59 49.93      3ICM
ATOM  51827 CD+3 CD3  159      93.607  32.867  -8.182  0.46 49.93      3ICM
ATOM  51828 CD+3 CD3  160     112.040  24.749  -2.624  0.33 49.93      3ICM
ATOM  51829 CD+3 CD3  161      91.969  45.413 -15.082  0.56 49.93      3ICM
ATOM  51830 CD+3 CD3  162      84.370  54.957 -18.909  0.39 49.93      3ICM
ATOM  51831 CD+3 CD3  163      89.554  59.874 -26.725  0.35 49.93      3ICM
ATOM  51832 CD+3 CD3  164      94.937  66.260 -28.025  0.32 49.93      3ICM
ATOM  51833 CD+3 CD3  165      98.691  71.798 -31.553  0.26 49.93      3ICM
ATOM  51834 CD+3 CD3  166     172.185  93.211  17.557  0.56 49.93      3ICM
ATOM  51835 CD+3 CD3  167     144.258  78.198  -7.629  0.73 49.93      3ICM
ATOM  51836 CD+3 CD3  168     126.932  88.530  -6.094  0.46 49.93      3ICM
ATOM  51837 CD+3 CD3  169     171.725  94.534   9.018  0.72 49.93      3ICM
ATOM  51838 CD+3 CD3  170     118.526  64.140  20.211  0.43 49.93      3ICM
ATOM  51839 CD+3 CD3  171     153.196 102.963  45.709  0.55 49.93      3ICM
ATOM  51840 CD+3 CD3  172     171.745  99.804  14.753  0.43 49.93      3ICM
ATOM  51841 CD+3 CD3  173     176.276  66.159   1.399  0.42 49.93      3ICM
ATOM  51842 CD+3 CD3  174     285.054 184.716 -47.604  0.38 49.93      3ICM
ATOM  51843 CD+3 CD3  175     291.977 188.836 -74.778  0.37 49.93      3ICM
ATOM  51844 CD+3 CD3  176     298.529  86.290 -46.164  0.44 49.93      3ICM
ATOM  51845 CD+3 CD3  177     177.768  91.757 -57.334  0.53 49.93      3ICM
ATOM  51846 CD+3 CD3  178     196.668  91.536 -40.732  0.78 49.93      3ICM
ATOM  51847 CD+3 CD3  179     178.446 112.017 -78.622  0.43 49.93      3ICM
ATOM  51848 CD+3 CD3  180     206.761  72.436  -8.196  0.68 49.93      3ICM
ATOM  51849 CD+3 CD3  181     272.609  98.153  11.653  0.66 49.93      3ICM
ATOM  51850 CD+3 CD3  182     295.898 110.456  12.297  0.27 49.93      3ICM
ATOM  51851 CD+3 CD3  183     202.758 119.706  11.288  0.49 49.93      3ICM
ATOM  51852 CD+3 CD3  184     198.747 130.245 -13.625  0.38 49.93      3ICM
ATOM  51853 CD+3 CD3  185     193.195 126.163 -16.338  0.09 49.93      3ICM
ATOM  51854 CD+3 CD3  186     255.175 124.260  -6.375  0.60 49.93      3ICM
ATOM  51855 CD+3 CD3  187     188.033 109.248 -29.708  0.44 49.93      3ICM
ATOM  51856 CD+3 CD3  188     186.976  96.260 -29.703  0.31 49.93      3ICM
END
```

CRYSTAL STRUCTURE OF THE 30S RIBOSOME AND ITS USE - FIGURE 2

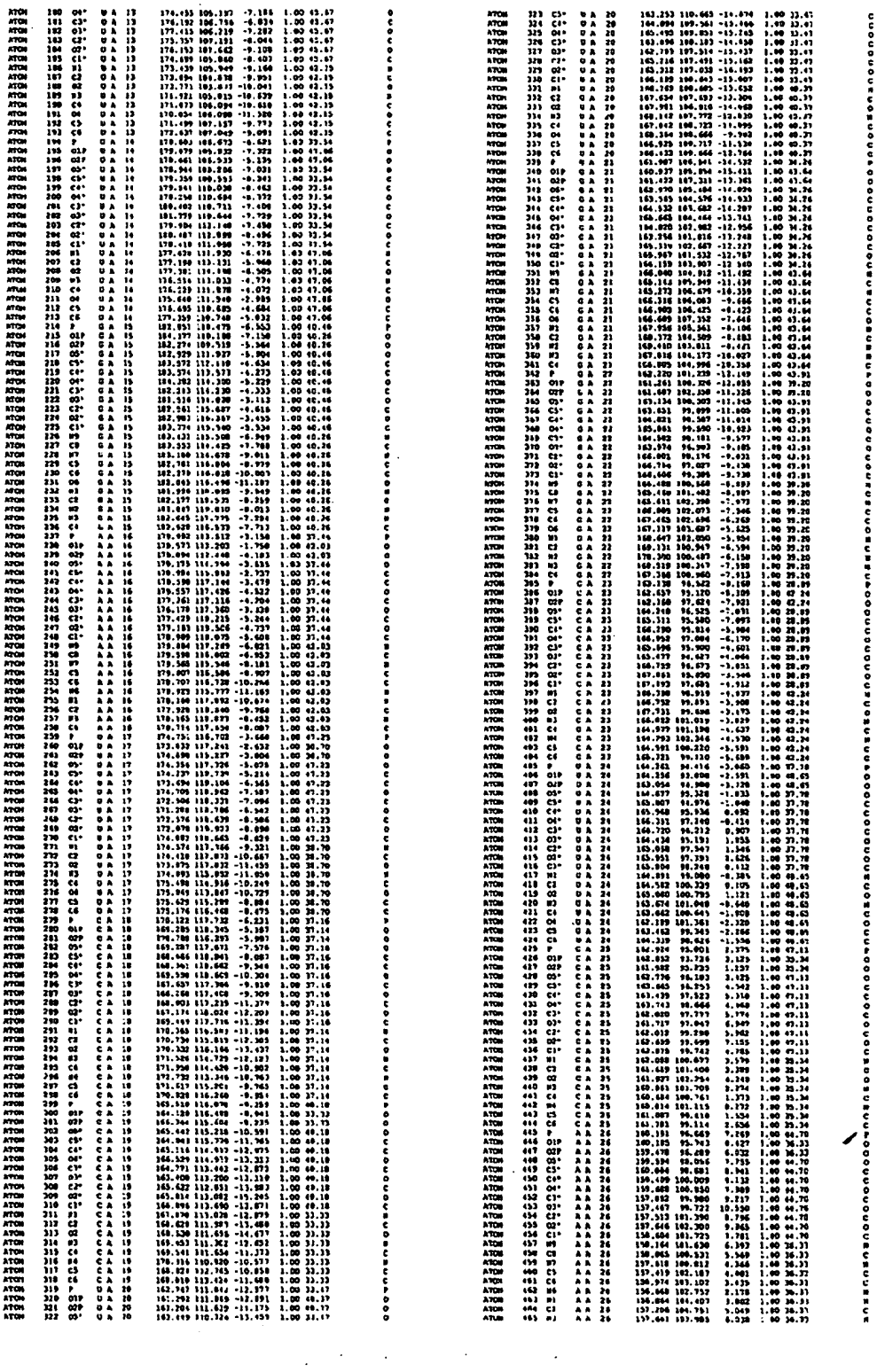
TABLE 1B

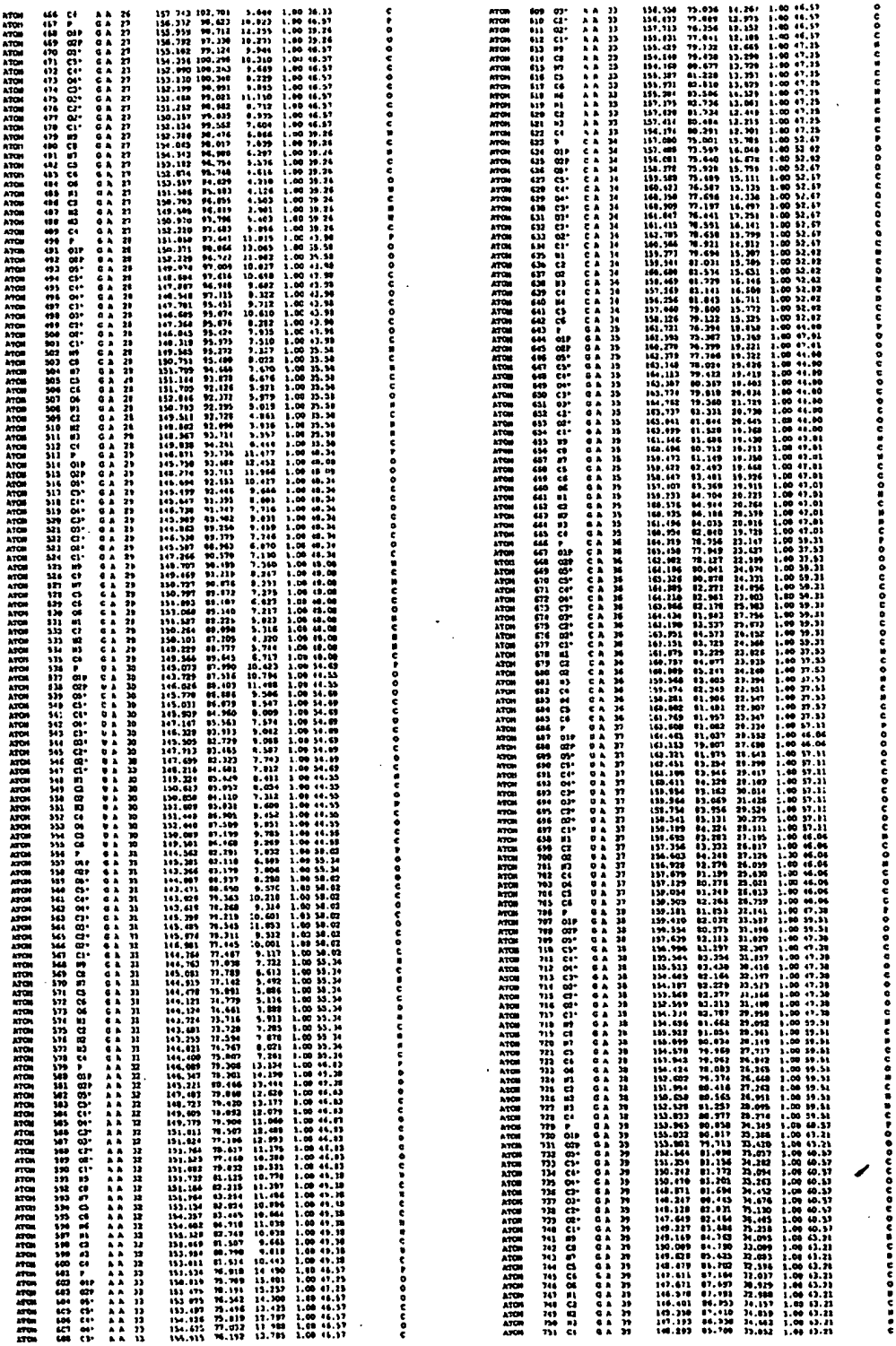
```
TITLE     STRUCTURE OF THE THERMUS THERMOPHILUS 30S RIBOSOMAL SUBUNIT
COMPND    MOL_ID: 1;
COMPND   2 MOLECULE: 16S RIBOSOMAL RNA;
COMPND   3 CHAIN: A;
COMPND   4 MOL_ID: 2;
COMPND   5 MOLECULE: FRAGMENT OF MESSENGER RNA;
COMPND   6 CHAIN: X;
COMPND   7 MOL_ID: 3;
COMPND   8 MOLECULE: 30S RIBOSOMAL PROTEIN S7;
COMPND   9 CHAIN: G;
COMPND  10 MOL_ID: 4;
COMPND  11 MOLECULE: 30S RIBOSOMAL PROTEIN S3;
COMPND  12 CHAIN: C;
COMPND  13 MOL_ID: 5;
COMPND  14 MOLECULE: 30S RIBOSOMAL PROTEIN S4;
COMPND  15 CHAIN: D;
COMPND  16 MOL_ID: 6;
COMPND  17 MOLECULE: 30S RIBOSOMAL PROTEIN S5;
COMPND  18 CHAIN: E;
COMPND  19 MOL_ID: 7;
COMPND  20 MOLECULE: 30S RIBOSOMAL PROTEIN S6;
COMPND  21 CHAIN: F;
COMPND  22 MOL_ID: 8;
COMPND  23 MOLECULE: 30S RIBOSOMAL PROTEIN S7;
COMPND  24 CHAIN: G;
COMPND  25 MOL_ID: 9;
COMPND  26 MOLECULE: 30S RIBOSOMAL PROTEIN S8;
COMPND  27 CHAIN: H;
COMPND  28 MOL_ID: 10;
COMPND  29 MOLECULE: 30S RIBOSOMAL PROTEIN S9;
COMPND  30 CHAIN: I;
COMPND  31 MOL_ID: 11;
COMPND  32 MOLECULE: 30S RIBOSOMAL PROTEIN S10;
COMPND  33 CHAIN: J;
COMPND  34 MOL_ID: 12;
COMPND  35 MOLECULE: 30S RIBOSOMAL PROTEIN S11;
COMPND  36 CHAIN: K;
COMPND  37 MOL_ID: 13;
COMPND  38 MOLECULE: 30S RIBOSOMAL PROTEIN S12;
COMPND  39 CHAIN: L;
COMPND  40 MOL_ID: 14;
COMPND  41 MOLECULE: 30S RIBOSOMAL PROTEIN S13;
COMPND  42 CHAIN: M;
COMPND  43 MOL_ID: 15;
COMPND  44 MOLECULE: 30S RIBOSOMAL PROTEIN S14;
COMPND  45 CHAIN: N;
COMPND  46 MOL_ID: 16;
COMPND  47 MOLECULE: 30S RIBOSOMAL PROTEIN S15;
COMPND  48 CHAIN: O;
COMPND  49 MOL_ID: 17;
COMPND  50 MOLECULE: 30S RIBOSOMAL PROTEIN S16;
COMPND  51 CHAIN: P;
COMPND  52 MOL_ID: 18;
COMPND  53 MOLECULE: 30S RIBOSOMAL PROTEIN S17;
COMPND  54 CHAIN: Q;
COMPND  55 MOL_ID: 19;
COMPND  56 MOLECULE: 30S RIBOSOMAL PROTEIN S18;
COMPND  57 CHAIN: R;
COMPND  58 MOL_ID: 20;
COMPND  59 MOLECULE: 30S RIBOSOMAL PROTEIN S19;
COMPND  60 CHAIN: S;
COMPND  61 MOL_ID: 21;
COMPND  62 MOLECULE: 30S RIBOSOMAL PROTEIN S20;
COMPND  63 CHAIN: T;
COMPND  64 MOL_ID: 22;
COMPND  65 MOLECULE: 30S RIBOSOMAL PROTEIN TH1;
COMPND  66 CHAIN: V
REMARK  103
REMARK  103 THE PROTEIN DATA BANK HAS ADOPTED THE SACCHARIDE CHEMISTS
REMARK  103 NOMENCLATURE FOR ATOMS OF THE DEOXYRIBOSE/RIBOSE MOIETY
REMARK  103 RATHER THAN THAT OF THE NUCLEOSIDE CHEMISTS.  THE RING
REMARK  103 OXYGEN ATOM IS LABELLED O4* INSTEAD OF O1*.

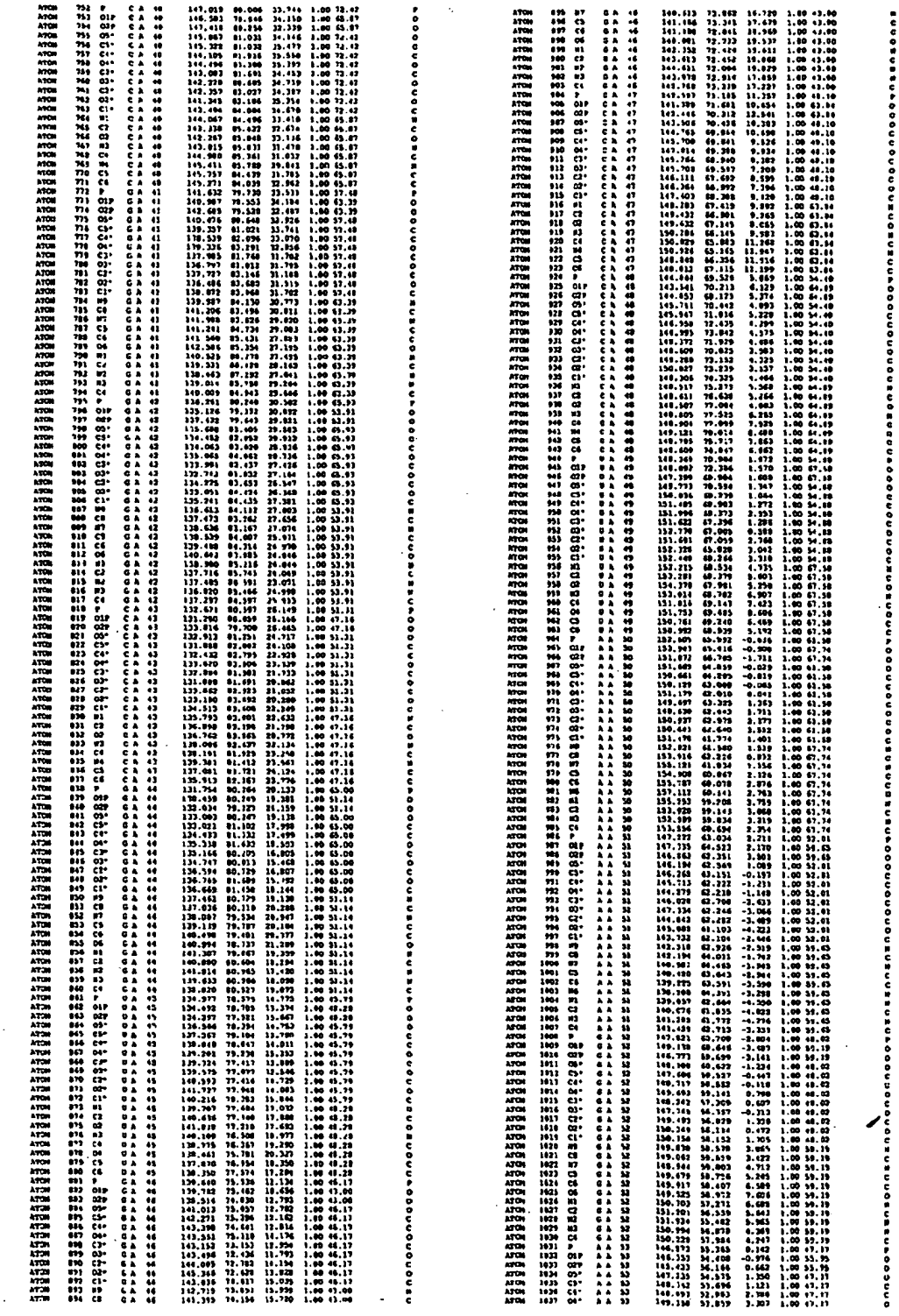
REMARK  290 CRYSTALLOGRAPHIC SYMMETRY
REMARK  290 SYMMETRY OPERATORS FOR SPACE GROUP: P 41 21 2
REMARK  290
REMARK  290      SYMOP   SYMMETRY
REMARK  290     NNNMM    OPERATOR
REMARK  290       1555   X,Y,Z
REMARK  290       2555   -X,-Y,1/2+Z
REMARK  290       3555   1/2-Y,1/2+X,1/4+Z
REMARK  290       4555   1/2+Y,1/2-X,3/4+Z
REMARK  290       5555   1/2-X,1/2+Y,1/4-Z
REMARK  290       6555   1/2+X,1/2-Y,3/4-Z
REMARK  290       7555   Y,X,-Z
REMARK  290       8555   -Y,-X,1/2-Z
REMARK  290
REMARK  290
REMARK  290      WHERE NNN -> OPERATOR NUMBER
REMARK  290            MMM -> TRANSLATION VECTOR

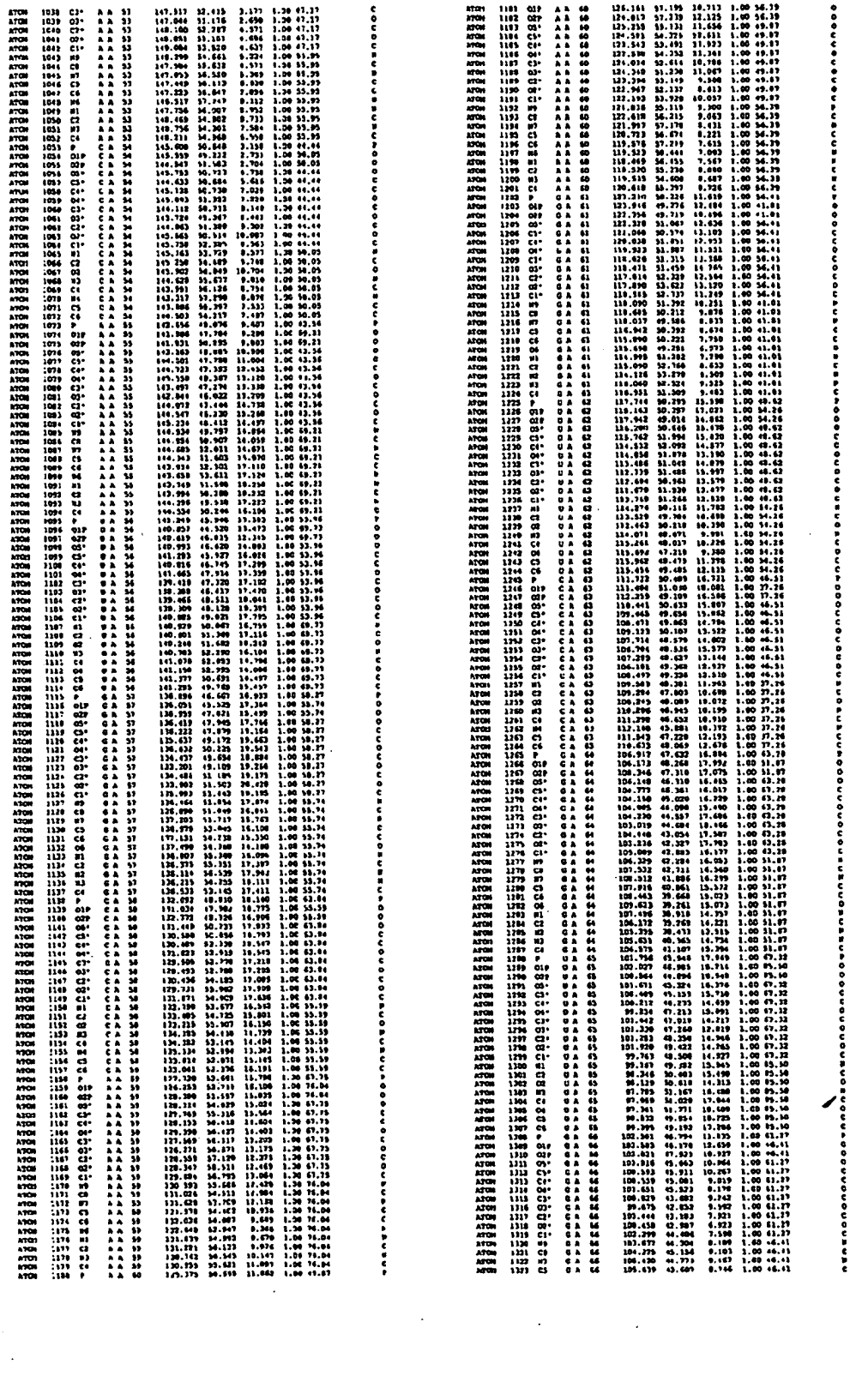
CRYST1  401.375  401.375  175.897  90.00  90.00  90.00 P 41 21 2     8

ORIGX1      1.000000  0.000000  0.000000        0.00000
ORIGX2      0.000000  1.000000  0.000000        0.00000
ORIGX3      0.000000  0.000000  1.000000        0.00000
SCALE1      0.002491  0.000000  0.000000        0.00000
SCALE2      0.000000  0.002491  0.000000        0.00000
SCALE3      0.000000  0.000000  0.005685        0.00000
ATOM      1  O5*  G A   3     133.566 110.584   1.562  1.00 52.26           O
ATOM      2  C5*  G A   3     133.754 109.301   2.183  1.00 52.26           C
ATOM      3  C4*  G A   3     134.524 109.369   3.516  1.00 52.26           C
ATOM      4  O4*  G A   3     134.871 111.130   3.944  1.00 57.24           O
ATOM      5  C3*  G A   3     134.992 108.949   4.918  1.00 52.26           C
ATOM      6  O3*  G A   3     135.593 107.662   4.935  1.00 52.26           O
ATOM      7  C2*  G A   3     135.321 109.677   5.798  1.00 52.26           C
ATOM      8  O2*  G A   3     137.119 109.752   5.858  1.00 52.26           O
ATOM      9  C1*  G A   3     135.473 111.268   5.713  1.00 52.26           C
ATOM     10  N1   G A   4     134.789 113.239   6.612  1.00 79.96           N
ATOM     11  C2   G A   4     133.983 112.396   7.367  1.00 79.96           C
ATOM     12  N2   G A   4     133.631 111.218   7.952  1.00 79.96           N
ATOM     13  N3   G A   4     134.095 113.452   8.812  1.00 79.96           N
ATOM     14  C4   G A   4     133.629 114.391   7.462  1.00 79.96           C
ATOM     15  O6   G A   4     132.848 115.344   8.165  1.00 79.96           O
ATOM     16  C5   G A   4     133.348 114.191   6.079  1.00 79.96           C
ATOM     17  C6   G A   4     133.910 113.179   5.278  1.00 79.96           C
ATOM     18  P    G A   4     134.020 107.045   6.138  1.00 49.51           P
ATOM     19  OIP  G A   4     136.784 105.589   6.153  1.00 58.37           O
ATOM     20  OIP  G A   4     136.434 107.567   6.886  1.00 58.37           O
ATOM     21  O5*  G A   4     134.035 107.997   4.739  1.00 49.51           O
ATOM     22  C5*  G A   4     132.948 107.029   3.478  1.00 49.51           C
ATOM     23  C4*  G A   4     132.443 107.369   2.727  1.00 49.51           C
ATOM     24  O4*  G A   4     130.737 109.043   2.179  1.00 49.51           O
ATOM     25  C3*  G A   4     131.686 108.613   1.496  1.00 49.51           C
ATOM     26  O3*  G A   4     131.589 107.589   0.509  1.00 49.51           O
ATOM     27  C2*  G A   4     131.080 109.739   2.495  1.00 49.51           C
ATOM     28  O2*  G A   4     131.852 109.223   1.594  1.00 49.51           O
ATOM     29  C1*  G A   4     131.698 109.638   3.896  1.00 49.51           C
ATOM     30  N9   G A   4     133.063 111.395   3.125  1.00 58.37           N
ATOM     31  C8   G A   4     132.518 112.239   1.895  1.00 58.37           C
ATOM     32  N7   G A   4     130.058 113.253   1.709  1.00 58.37           N
ATOM     33  C5   G A   4     130.326 112.215   2.586  1.00 58.37           C
ATOM     34  C6   G A   4     129.067 110 110   4.123  1.00 58.37           C
ATOM     35  O6   G A   4     127.329 115.255   4.673  1.00 58.37           O
ATOM     36  N1   G A   4     128.947 113.637   5.341  1.00 58.37           N
ATOM     37  C2   G A   4     129.109 112.436   5.516  1.00 58.37           C
ATOM     38  N2   G A   4     127.982 114.130   6.784  1.00 58.37           N
ATOM     39  N3   G A   4     130.351 111.548   6.538  1.00 58.37           N
ATOM     40  C4   G A   4     130.939 112.075   5.345  1.00 58.37           C
ATOM     41  P    G A   7     143.078 108.145   6.245  1.00 47.85           P
ATOM     42  OIP  G A   7     141.648 106.998   6.981  1.00 56.02           O
ATOM     43  OIP  G A   7     142.999 109.442   6.939  1.00 56.02           O
ATOM     44  O5*  G A   7     143.992 108.427   3.939  1.00 47.85           O
ATOM     45  C5*  G A   7     144.954 107.478   3.518  1.00 47.85           C
ATOM     46  C4*  G A   7     145.331 108.153   2.593  1.00 47.85           C
ATOM     47  O4*  G A   7     145.298 109.435   3.260  1.00 47.85           O
ATOM     48  C3*  G A   7     144.409 109.495   2.109  1.00 47.85           C
ATOM     49  O3*  G A   7     143.578 109.417   2.939  1.00 47.85           O
ATOM     50  C2*  G A   7     145.197 110.451   1.217  1.00 47.85           C
ATOM     51  O2*  G A   7     147.637 111.044   0.639  1.00 47.85           O
ATOM     52  C1*  G A   7     145.963 109.346  -0.234  1.00 47.85           C
ATOM     53  N9   G A   7     145.647 110.192  -1.135  1.00 50.02           N
ATOM     54  C8   G A   7     145.779 109.632  -0.093  1.00 50.02           C
ATOM     55  N7   G A   7     142.182 111.144  -1.936  1.00 50.02           N
ATOM     56  C5   G A   7     144.138 111.863  -2.944  1.00 50.02           C
ATOM     57  C6   G A   7     144.972 111.663  -4.030  1.00 50.02           C
ATOM     58  O6   G A   7     146.144 112.049  -4.083  1.00 50.02           O
ATOM     59  N1   G A   7     144.236 111.179  -4.978  1.00 50.02           N
ATOM     60  C2   G A   7     142.858 110.548  -4.603  1.00 50.02           C
ATOM     61  N2   G A   7     142.388 110.389  -5.647  1.00 50.02           N
ATOM     62  N3   G A   7     142.472 110.187  -3.379  1.00 50.02           N
ATOM     63  C4   G A   7     143.283 110.451  -2.479  1.00 50.02           C
ATOM     64  P    A A   8     148.942 108.716   3.388  1.00 48.35           P
ATOM     65  OIP  A A   8     148.899 106.776   0.901  1.00 46.78           O
ATOM     66  OIP  A A   8     149.139 107.317   3.145  1.00 46.78           O
ATOM     67  O5*  A A   8     150.994 109.003   7.889  1.00 48.35           O
ATOM     68  C5*  A A   8     150.261 111.000   9.212  1.00 48.35           C
ATOM     69  C4*  A A   8     151.671 111.562   9.137  1.00 48.35           C
ATOM     70  O4*  A A   8     151.796 111.998   9.792  1.00 48.35           O
ATOM     71  C3*  A A   8     152.727 110.416   9.261  1.00 48.35           C
ATOM     72  O3*  A A   8     153.315 110.906   1.717  1.00 48.35           O
ATOM     73  C2*  A A   8     152.999 109.974   9.683  1.00 48.35           C
ATOM     74  O2*  A A   8     154.291 109.452   9.892  1.00 48.35           O
ATOM     75  C1*  A A   8     152.907 111.282   9.457  1.00 48.35           C
ATOM     76  N9   A A   8     153.909 111.089   8.438  1.00 46.78           N
ATOM     77  C8   A A   8     153.122 110.920   6.991  1.00 46.78           C
ATOM     78  N7   A A   8     153.000 110.792   7.621  1.00 46.78           N
ATOM     79  C5   A A   8     153.416 110.861   8.814  1.00 46.78           C
ATOM     80  C6   A A   8     153.035 110.772   9.966  1.00 46.78           C
ATOM     81  N6   A A   8     152.369 110.597  10.901  1.00 46.78           N
ATOM     82  N1   A A   8     154.289 110.871   9.908  1.00 46.78           N
ATOM     83  C2   A A   8     155.024 111.051   8.837  1.00 46.78           C
ATOM     84  N3   A A   8     154.559 111.152   6.923  1.00 46.78           N
ATOM     85  C4   A A   8     153.293 111.046   6.522  1.00 46.78           C
ATOM     86  P    A A   9     154.163 110.948  -0.129  1.00 35.74           P
ATOM     87  OIP  A A   9     153.617 112.241  -0.424  1.00 38.06           O
ATOM     88  OIP  A A   9     153.678 109.660  -0.597  1.00 38.06           O
ATOM     89  O5*  A A   9     156.747 110.945  -0.809  1.00 35.74           O
ATOM     90  C5*  A A   9     156.551 111.899   0.785  1.00 35.74           C
ATOM     91  C4*  A A   9     157.207 111.033   1.901  1.00 35.74           C
ATOM     92  O4*  A A   9     156.906 109.992   2.149  1.00 35.74           O
ATOM     93  C3*  A A   9     158.732 110.988   1.679  1.00 35.74           C
ATOM     94  O3*  A A   9     159.513 111.660   1.908  1.00 35.74           O
ATOM     95  C2*  A A   9     158.787 109.888   2.719  1.00 35.74           C
ATOM     96  O2*  A A   9     158.965 109.992   4.026  1.00 35.74           O
ATOM     97  C1*  A A   9     157.417 108.839   2.672  1.00 35.74           C
ATOM     98  N9   A A   9     157.456 107.689   1.769  1.00 38.06           N
ATOM     99  C8   A A   9     156.781 107.537   0.581  1.00 38.06           C
ATOM    100  N7   A A   9     157.881 106.629  -0.033  1.00 38.06           N
ATOM    101  C5   A A   9     158.795 107.006   0.906  1.00 38.06           C
ATOM    102  C6   A A   9     159.965 106.579   0.670  1.00 38.06           C
ATOM    103  N6   A A   9     160.535 105.789  -0.261  1.00 38.06           N
ATOM    104  N1   A A   9     159.467 106.990   1.766  1.00 38.06           N
ATOM    105  C2   A A   9     158.861 107.153   2.946  1.00 38.06           C
ATOM    106  N3   A A   9     158.475 106.795   3.797  1.00 38.06           N
ATOM    107  C4   A A   9     159.020 106.292   2.673  1.00 38.06           C
ATOM    108  P    A A  10     160.790 111.721   1.395  1.00 38.12           P
ATOM    109  OIP  A A  10     161.460 113.000   1.241  1.00 41.31           O
ATOM    110  OIP  A A  10     160.658 111.352  -0.261  1.00 41.31           O
ATOM    111  O5*  A A  10     161.588 110.719   2.541  1.00 38.12           O
ATOM    112  C5*  A A  10     163.865 110.964   3.324  1.00 38.12           C
ATOM    113  C4*  A A  10     163.986 109.845   4.342  1.00 38.12           C
ATOM    114  O4*  A A  10     163.979 108.616   3.973  1.00 38.12           O
ATOM    115  C3*  A A  10     162.893 109.793   5.367  1.00 38.12           C
ATOM    116  O3*  A A  10     163.286 110.170   6.663  1.00 38.12           O
ATOM    117  C2*  A A  10     164.337 107.948   3.708  1.00 38.12           C
ATOM    118  O2*  A A  10     164.692 107.868   5.059  1.00 38.12           O
ATOM    119  C1*  A A  10     162.787 107.482   3.621  1.00 38.12           C
ATOM    120  N9   A A  10     162.438 106.986   2.288  1.00 41.31           N
ATOM    121  C8   A A  10     161.746 107.636   1.371  1.00 41.31           C
ATOM    122  N7   A A  10     161.549 106.958   0.223  1.00 41.31           N
ATOM    123  C5   A A  10     162.168 105.758   0.491  1.00 41.31           C
ATOM    124  C6   A A  10     162.307 104.586  -0.264  1.00 41.31           C
ATOM    125  N6   A A  10     161.826 104.449  -1.502  1.00 41.31           N
ATOM    126  N1   A A  10     162.968 103.549   0.298  1.00 41.31           N
ATOM    127  C2   A A  10     163.447 103.672   1.531  1.00 41.31           C
ATOM    128  N3   A A  10     163.388 104.743   2.344  1.00 41.31           N
ATOM    129  C4   A A  10     162.723 105.732   1.757  1.00 41.31           C
ATOM    130  P    A A  11     165.498 110.361   2.643  1.00 33.90           P
ATOM    131  OIP  A A  11     166.346 111.362   3.284  1.00 33.90           O
ATOM    132  OIP  A A  11     165.903 110.543   1.267  1.00 33.90           O
ATOM    133  O5*  A A  11     167.132 108.949   2.888  1.00 40.23           O
ATOM    134  C5*  A A  11     167.746 108.498   3.930  1.00 40.23           C
ATOM    135  C4*  A A  11     168.428 107.189   3.698  1.00 40.23           C
ATOM    136  O4*  A A  11     167.437 106.176   3.438  1.00 40.23           O
ATOM    137  C3*  A A  11     169.363 107.130   2.498  1.00 40.23           C
ATOM    138  O3*  A A  11     169.365 107.120   1.098  1.00 40.23           O
ATOM    139  C2*  A A  11     170.733 107.064   3.164  1.00 40.23           C
ATOM    140  O2*  A A  11     171.644 107.668   2.284  1.00 40.23           O
ATOM    141  C1*  A A  11     170.926 105.758   3.975  1.00 40.23           C
ATOM    142  N9   A A  11     171.232 104.572   3.043  1.00 40.23           N
ATOM    143  C8   A A  11     171.976 105.361   2.275  1.00 40.23           C
ATOM    144  N7   A A  11     172.154 104.871   1.052  1.00 40.23           N
ATOM    145  C5   A A  11     171.526 103.704   1.265  1.00 40.23           C
ATOM    146  C6   A A  11     171.230 102.615   0.465  1.00 40.23           C
ATOM    147  N6   A A  11     171.619 102.477  -0.753  1.00 40.23           N
ATOM    148  N1   A A  11     170.553 101.630   1.060  1.00 40.23           N
ATOM    149  C2   A A  11     170.164 101.768   2.324  1.00 40.23           C
ATOM    150  N3   A A  11     170.393 102.796   3.141  1.00 40.23           N
ATOM    151  C4   A A  11     171.026 103.843   2.579  1.00 40.23           C
ATOM    152  P    A A  12     171.516 104.544  -0.170  1.00 41.73           P
ATOM    153  OIP  A A  12     171.633 105.710  -1.116  1.00 41.73           O
ATOM    154  OIP  A A  12     171.126 103.268  -0.765  1.00 41.73           O
ATOM    155  O5*  A A  12     172.888 104.358   0.717  1.00 41.73           O
ATOM    156  C5*  A A  12     173.742 106.293   0.959  1.00 41.73           C
ATOM    157  C4*  A A  12     173.792 106.257  -0.581  1.00 41.73           C
ATOM    158  O4*  A A  12     171.515 104.544  -0.178  1.00 41.73           O
ATOM    159  C3*  A A  12     173.633 105.710  -1.516  1.00 41.73           C
ATOM    160  O3*  A A  12     174.109 105.958  -2.817  1.00 41.73           O
ATOM    161  C2*  A A  12     174.488 104.558  -2.177  1.00 41.73           C
ATOM    162  O2*  A A  12     173.280 103.478  -2.468  1.00 41.73           O
ATOM    163  C1*  A A  12     172.264 104.145  -3.519  1.00 41.73           C
ATOM    164  N9   A A  12     170.861 104.810  -3.279  1.00 48.69           N
ATOM    165  C8   A A  12     169.723 104.129  -2.879  1.00 48.69           C
ATOM    166  N7   A A  12     168.698 104.963  -3.193  1.00 48.69           N
ATOM    167  C5   A A  12     169.182 106.181  -3.568  1.00 48.69           C
ATOM    168  C6   A A  12     168.173 104.851  -3.368  1.00 48.69           C
ATOM    169  N6   A A  12     167.348 106.127  -3.962  1.00 48.69           N
ATOM    170  N1   A A  12     168.948 104.778  -2.679  1.00 48.69           N
ATOM    171  C2   A A  12     169.686 106.719  -2.064  1.00 48.69           C
ATOM    172  N3   A A  12     170.547 107.378  -1.581  1.00 48.69           N
ATOM    173  C4   A A  12     170.043 107.236  -2.593  1.00 48.69           C
ATOM    174  P    A A  13     174.043 107.978  -3.884  1.00 42.15           P
ATOM    175  OIP  A A  13     173.608 108.017  -2.908  1.00 42.15           O
ATOM    176  OIP  A A  13     175.423 106.673  -3.956  1.00 42.15           O
ATOM    177  O5*  A A  13     174.345 108.387  -4.983  1.00 45.67           O
ATOM    178  C5*  A A  13     175.348 107.917  -4.168  1.00 45.67           C
```

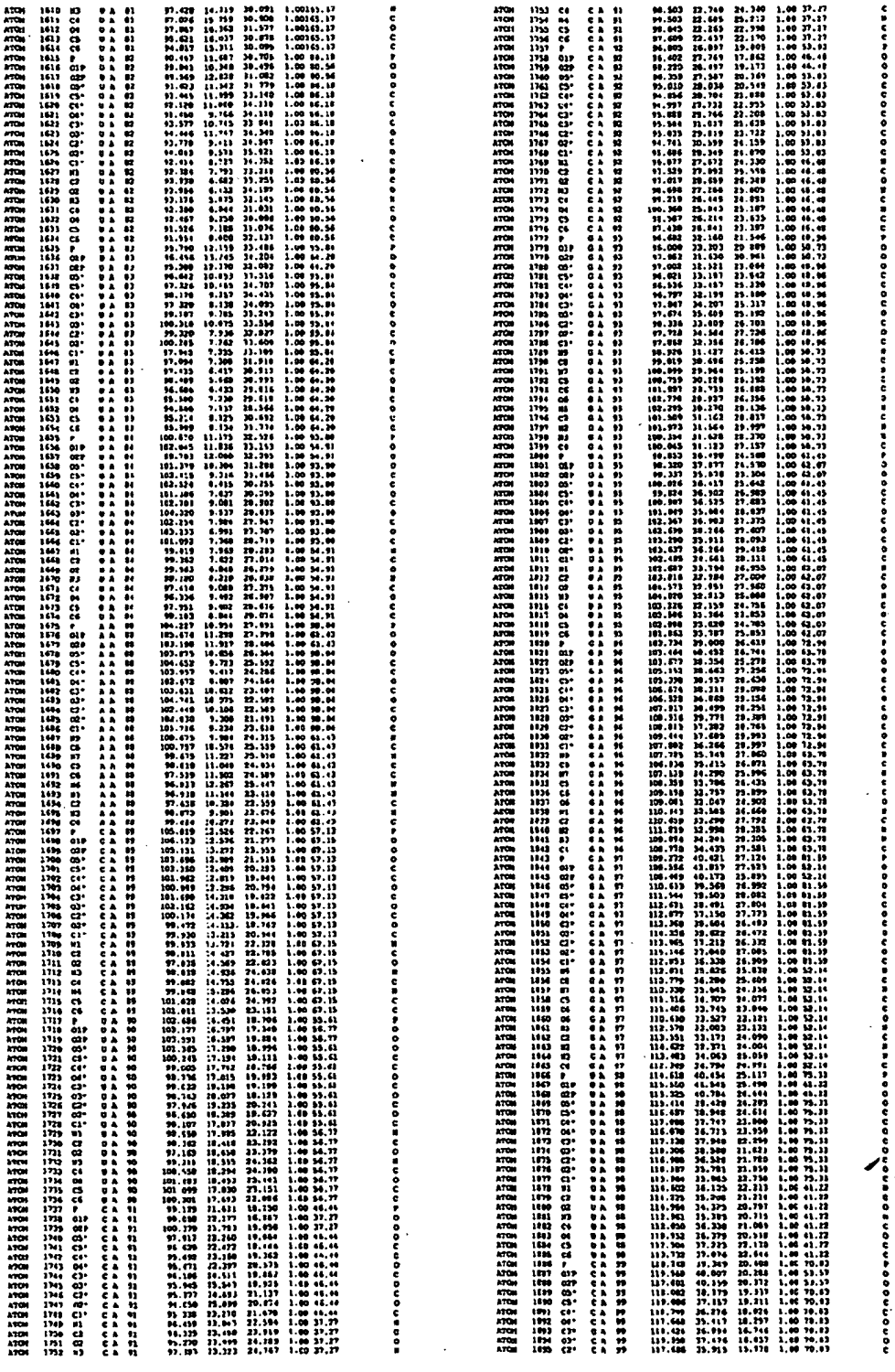


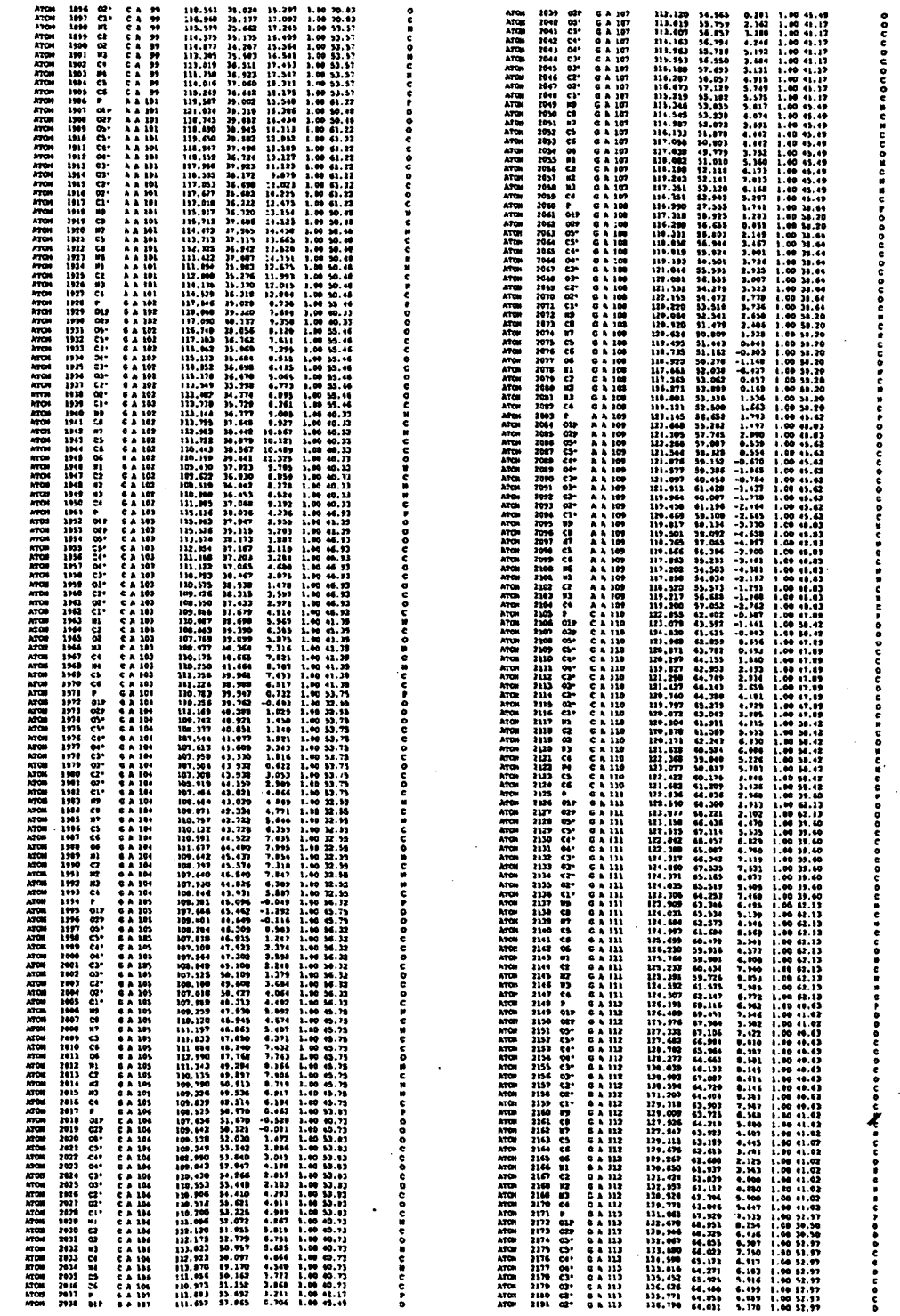


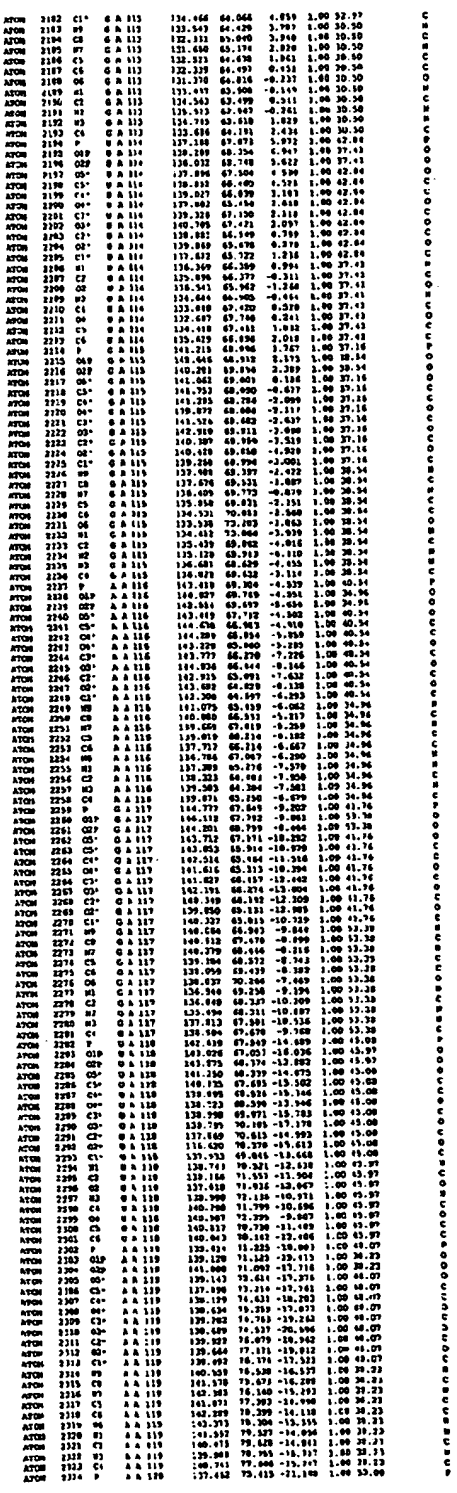


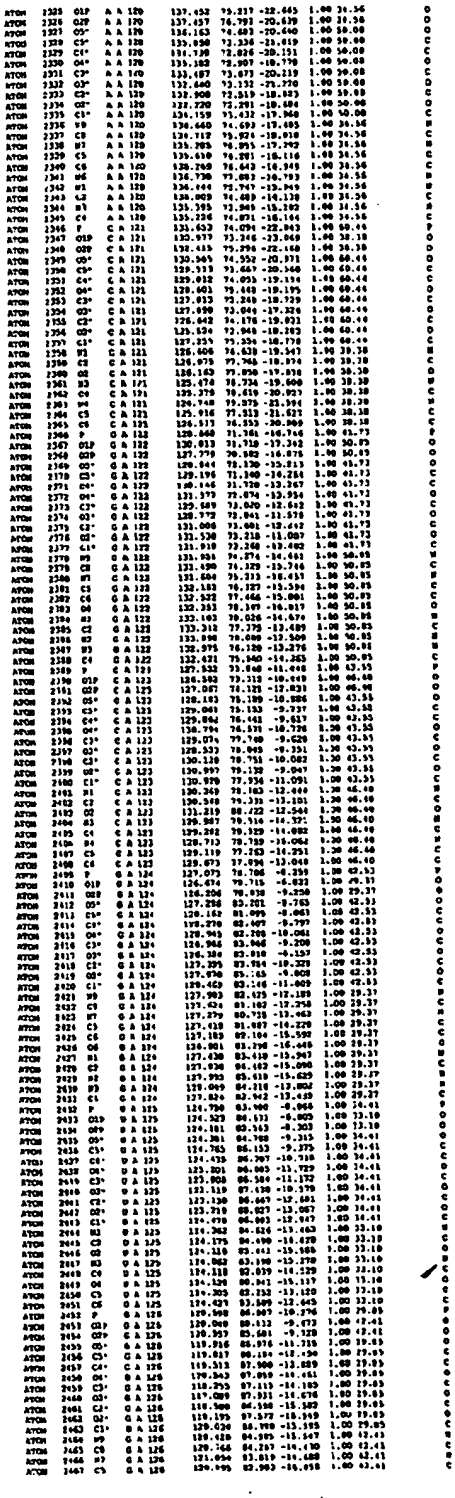


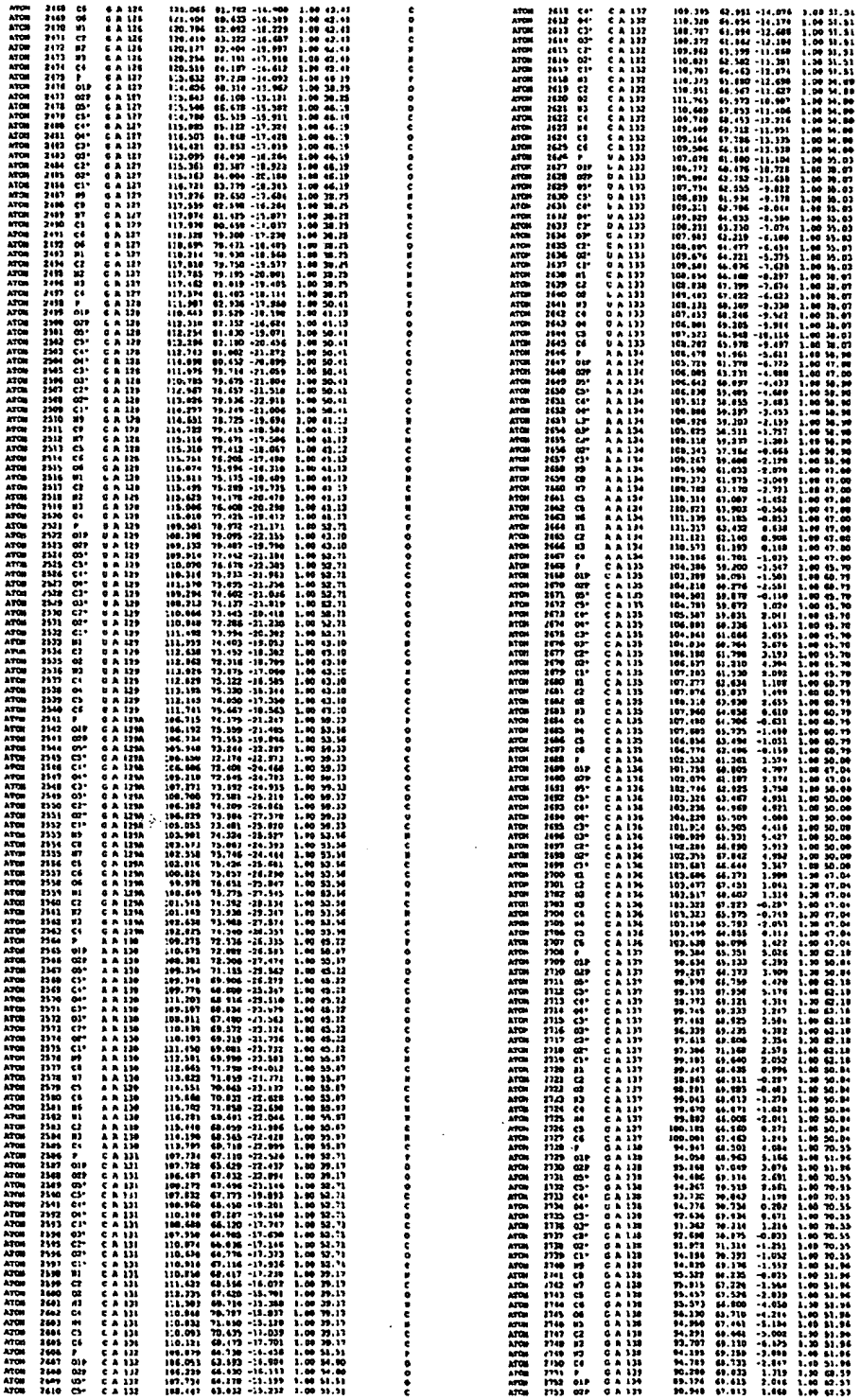


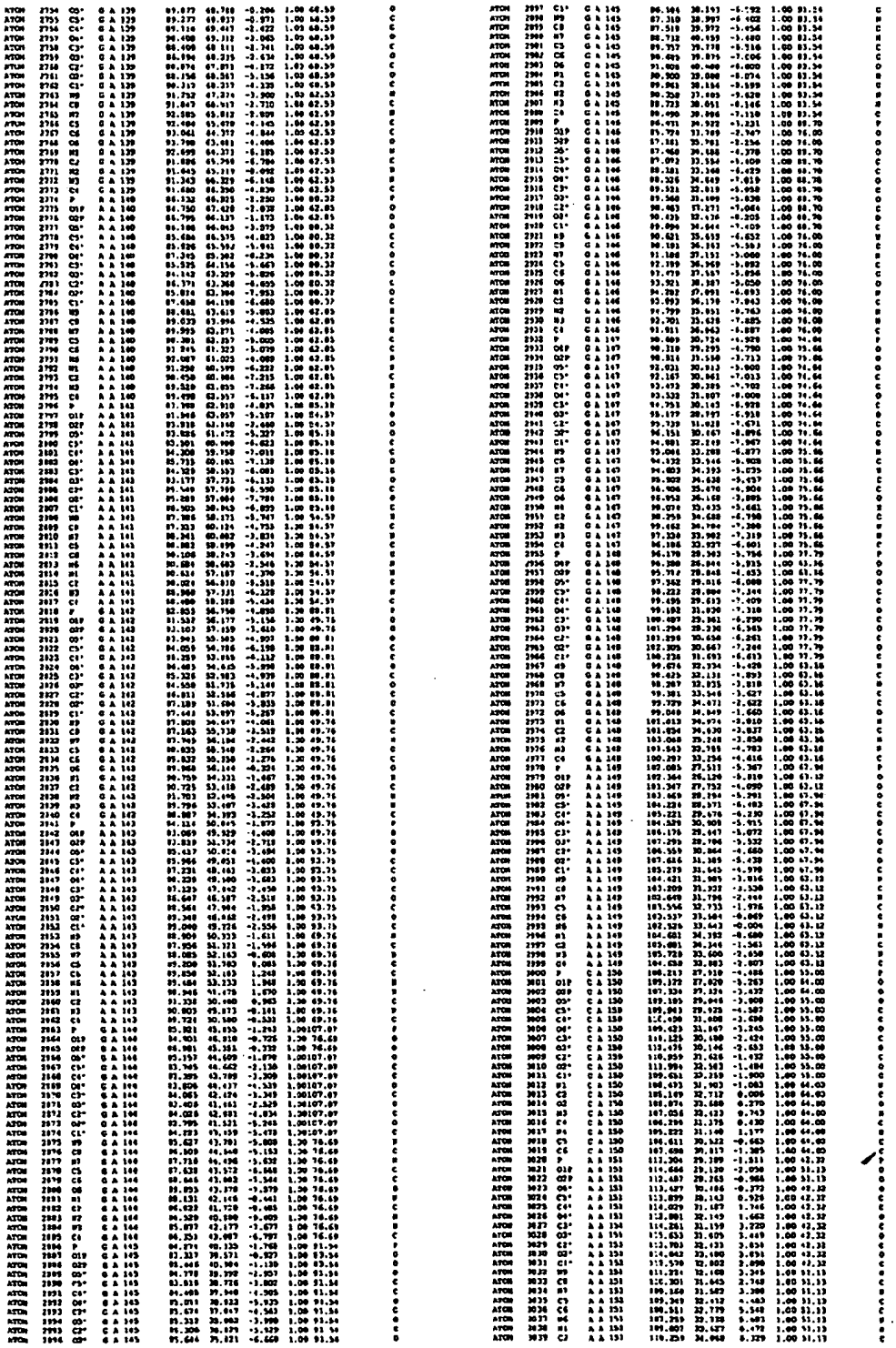


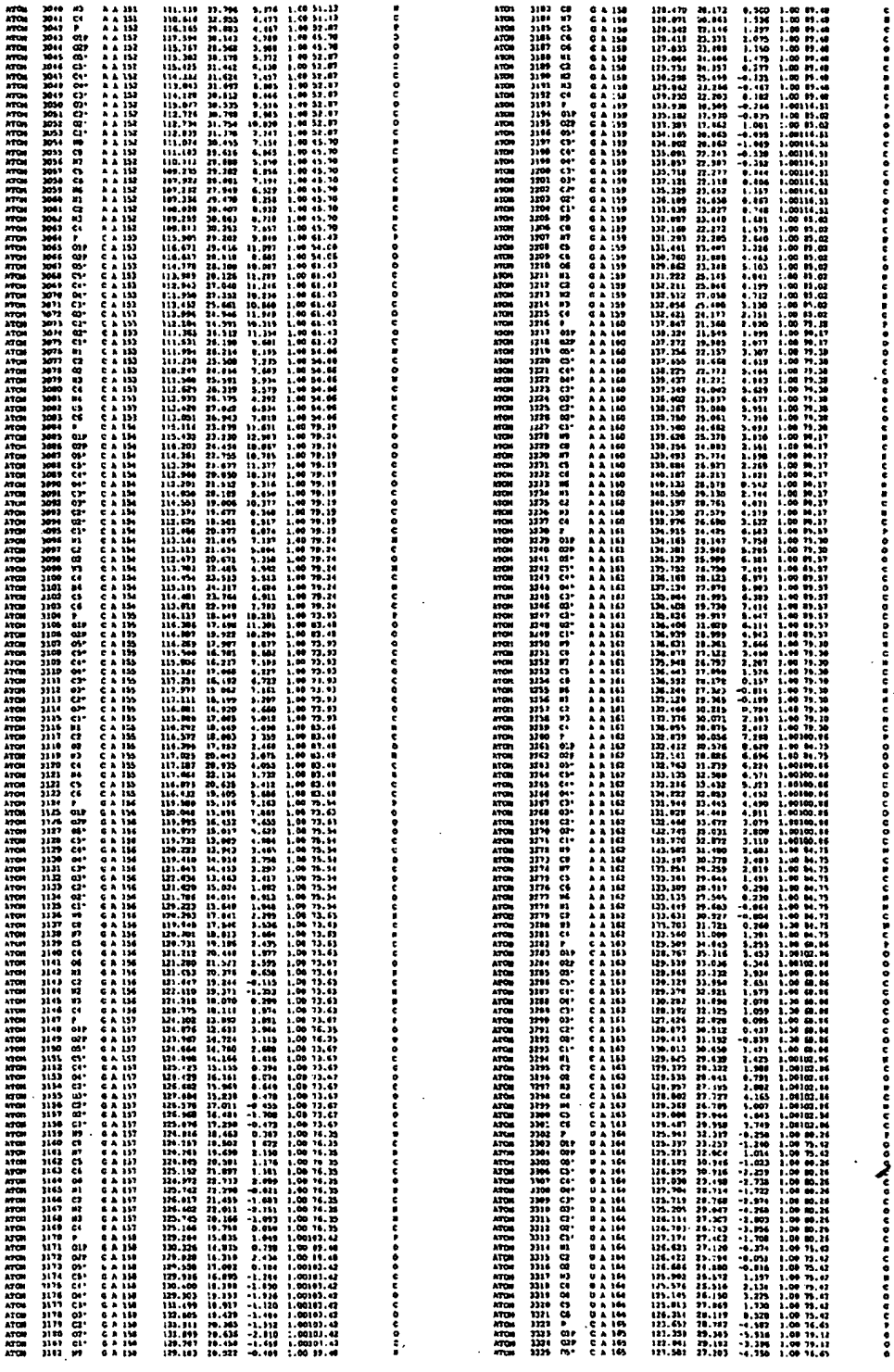


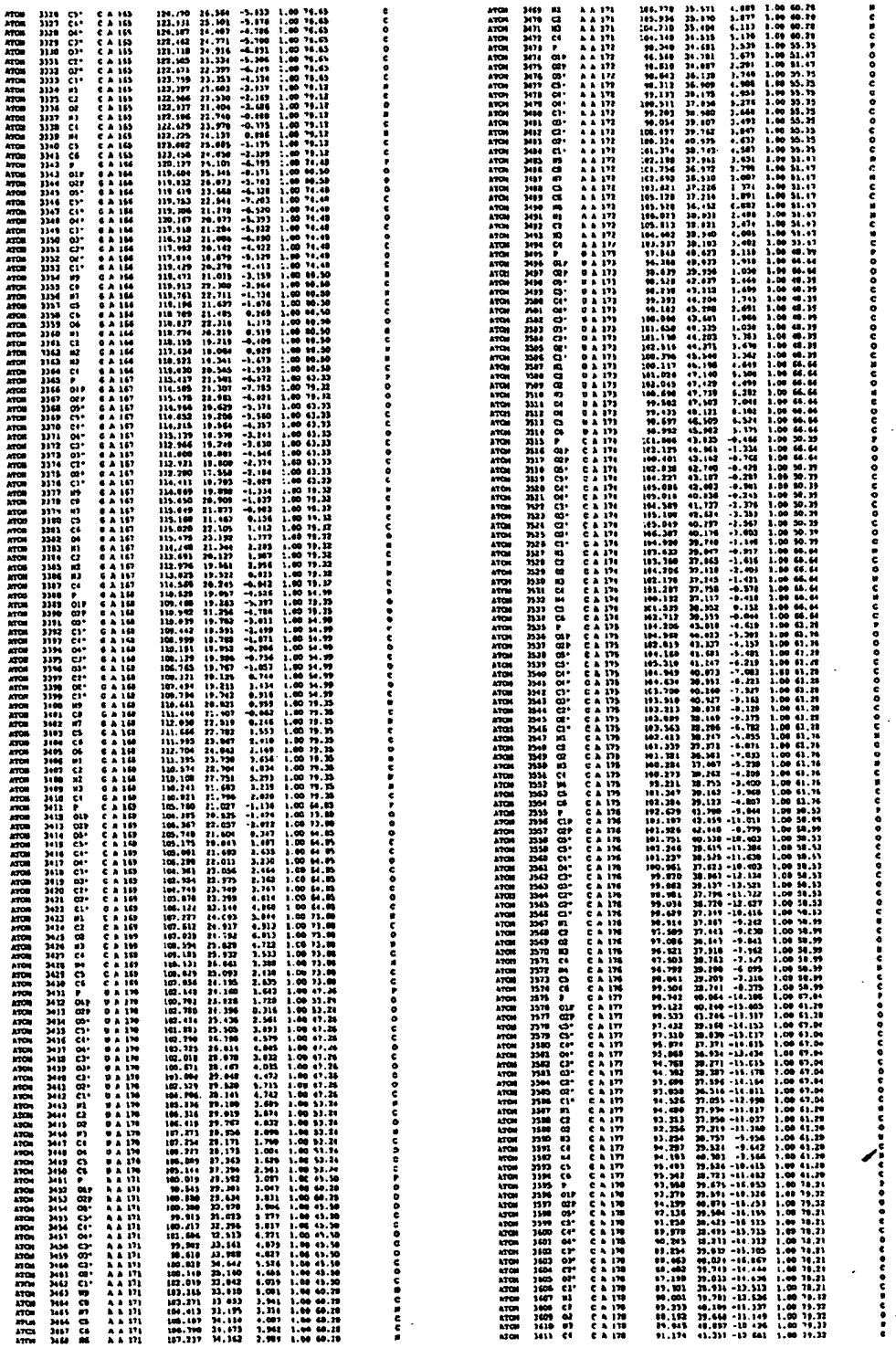


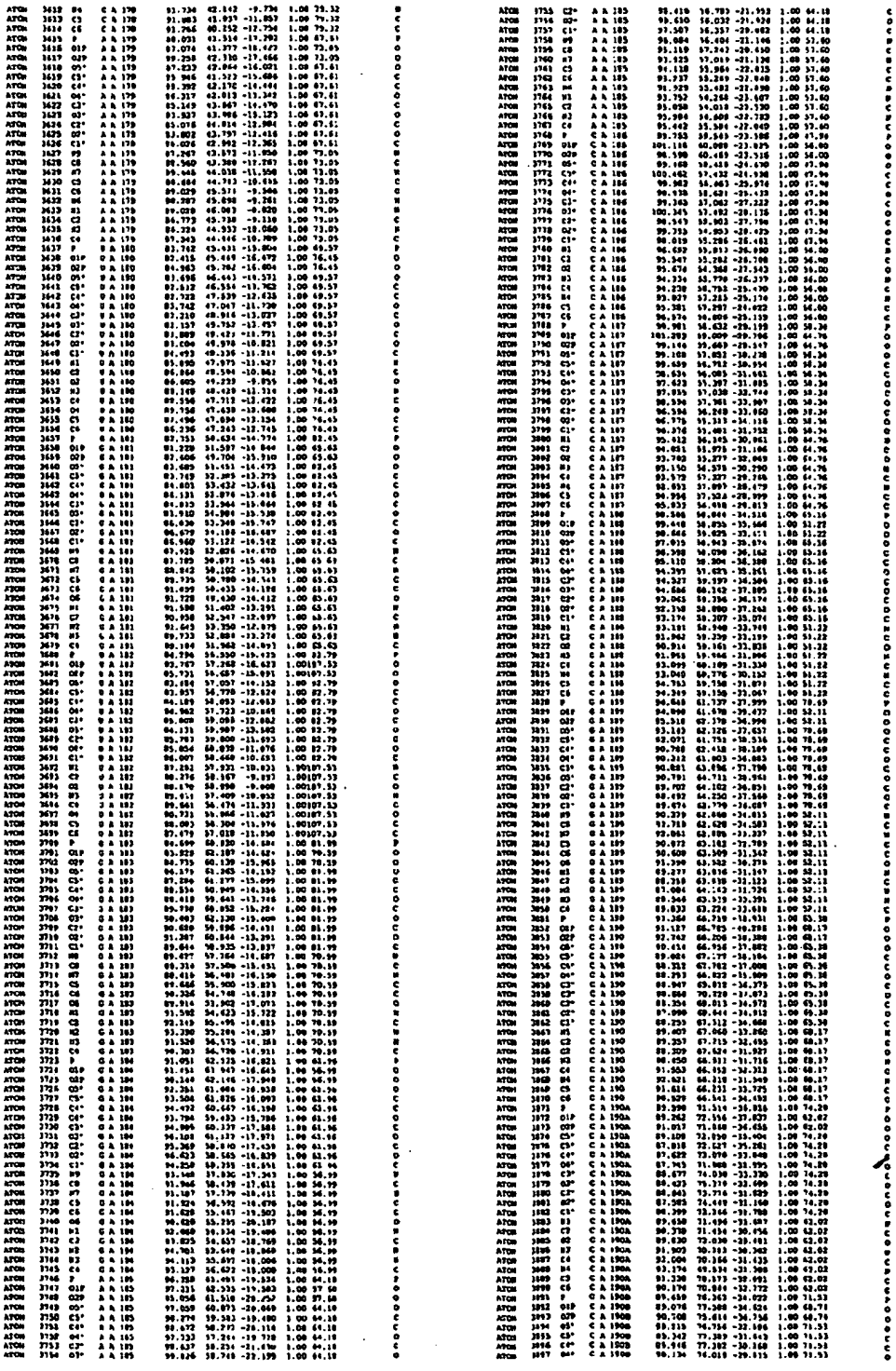


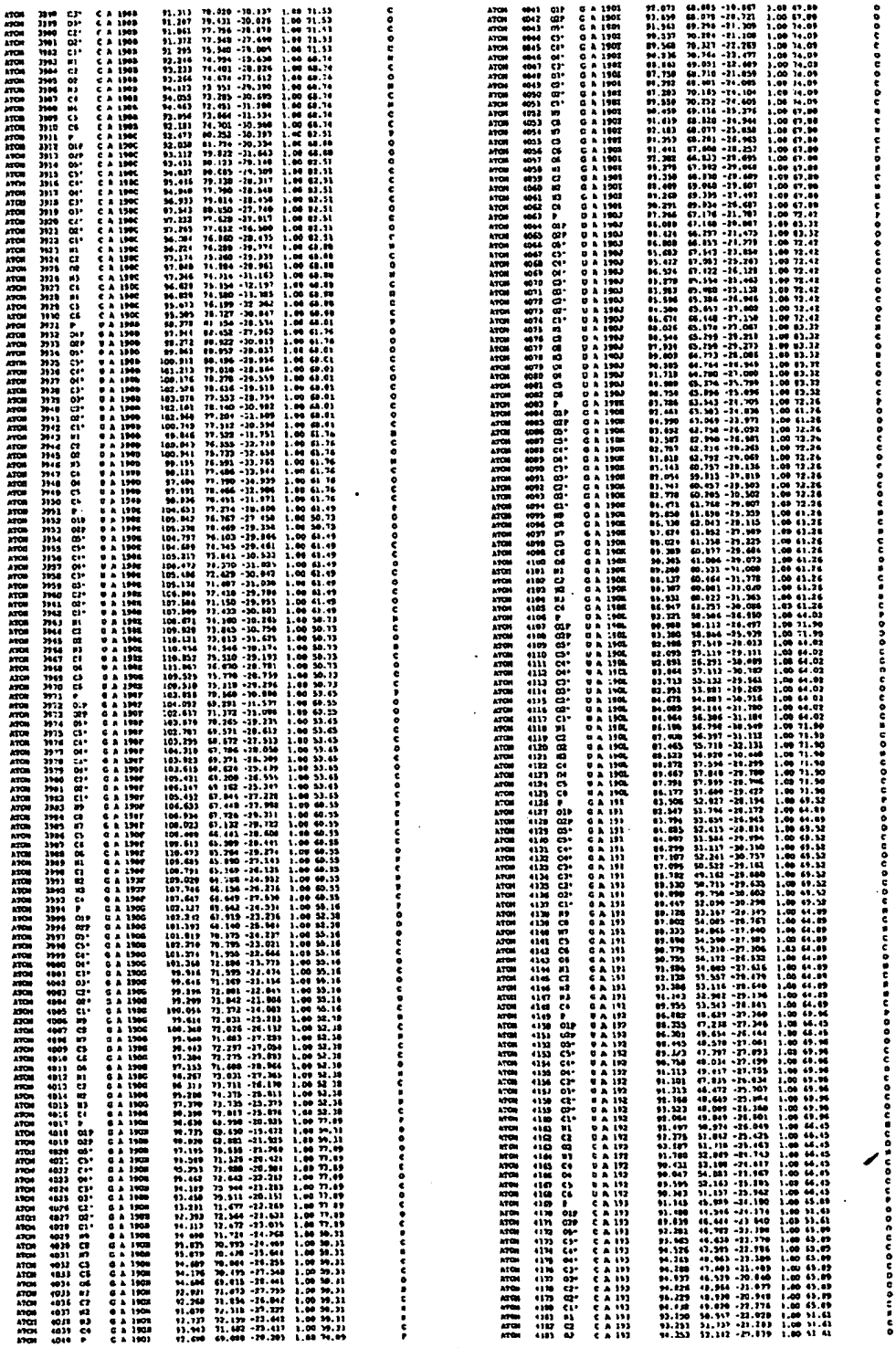


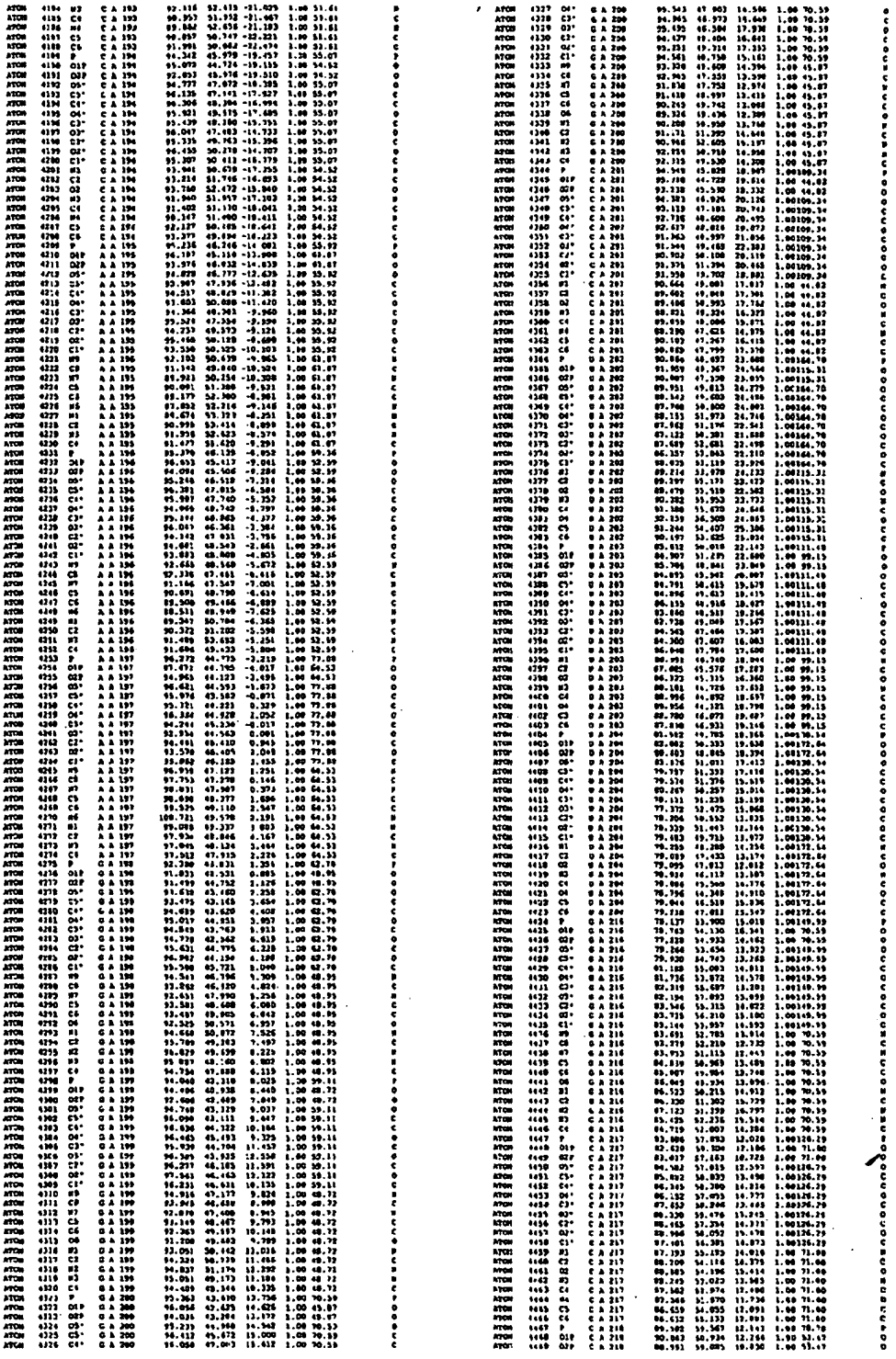


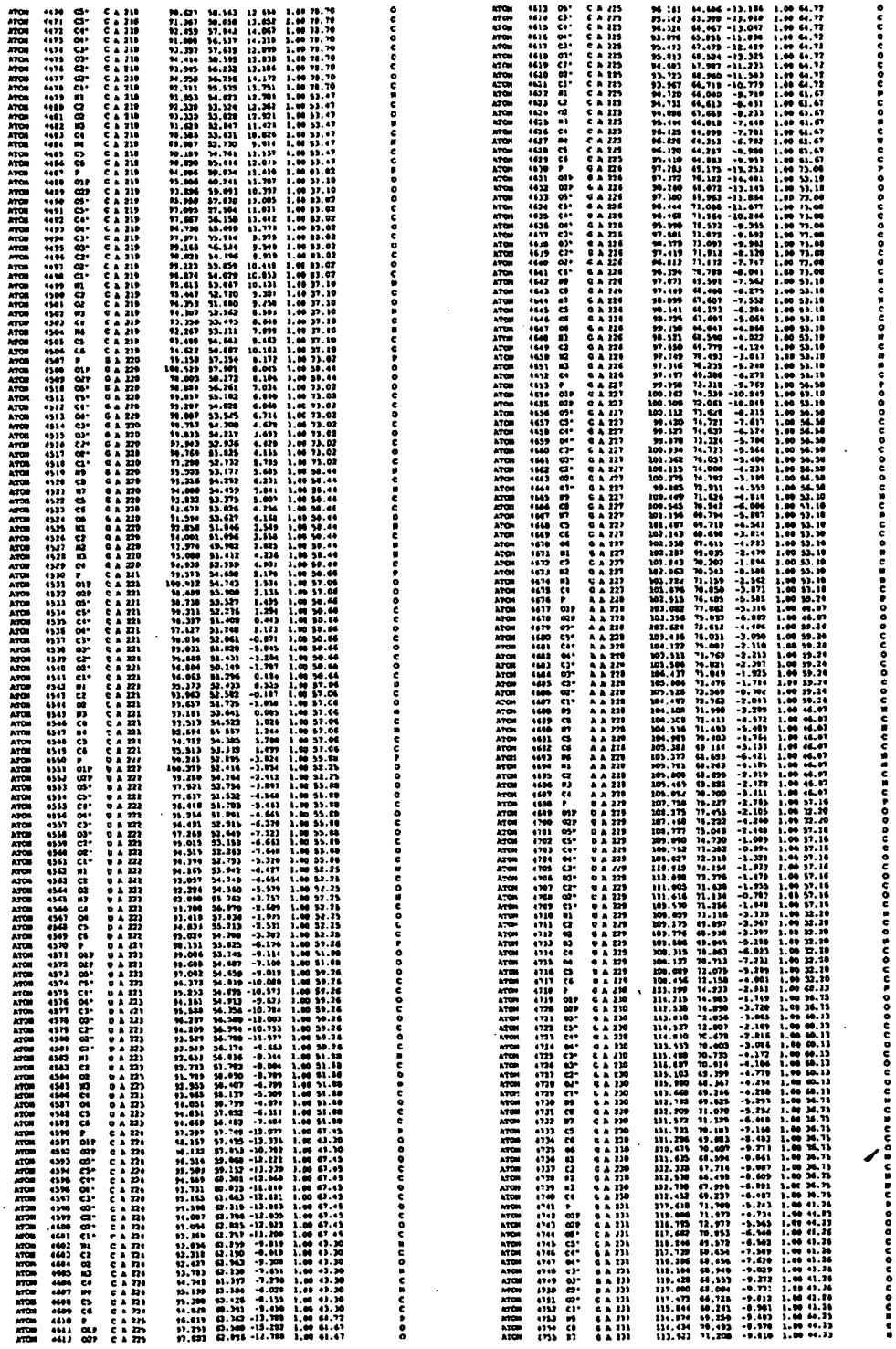


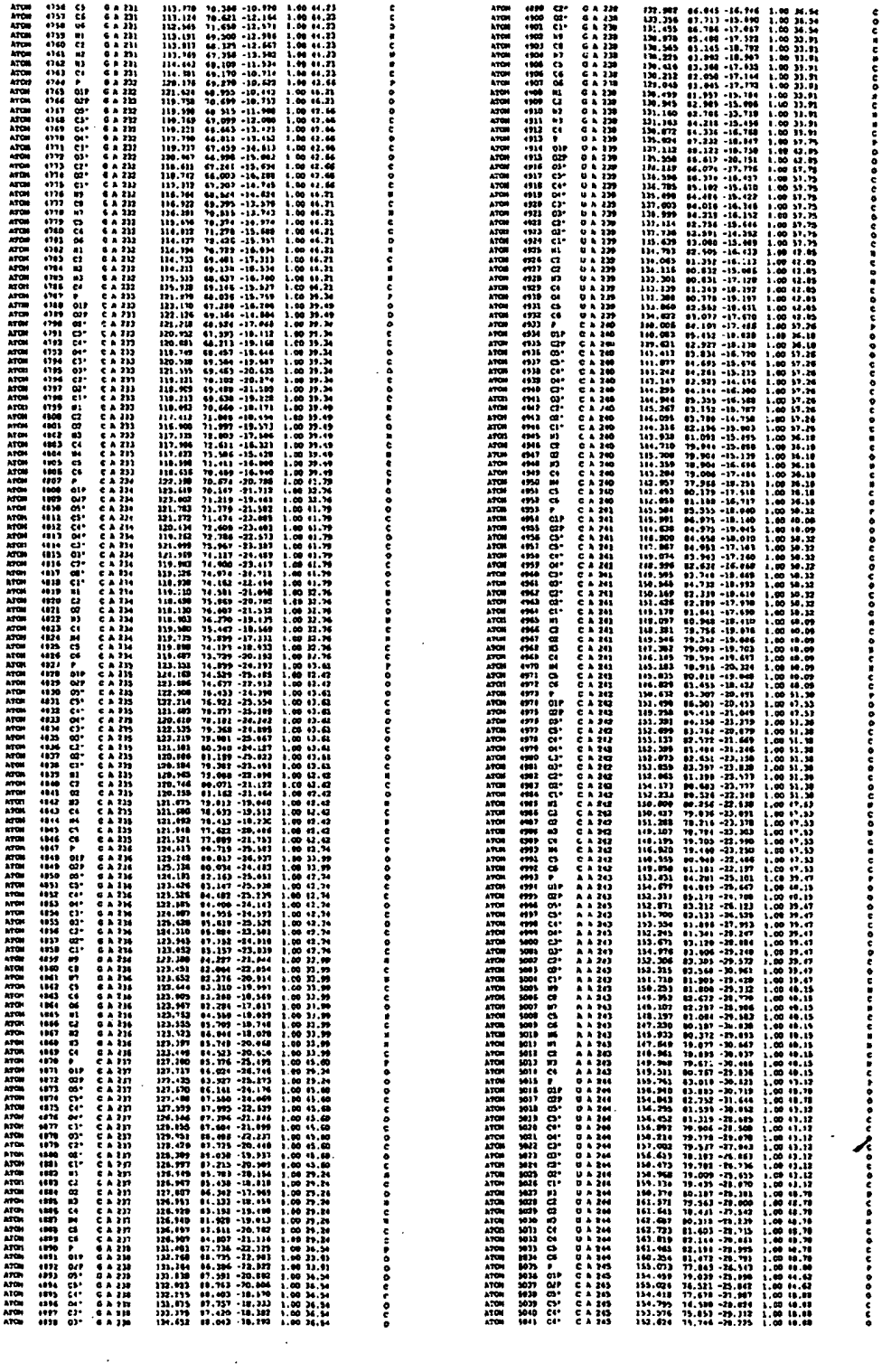


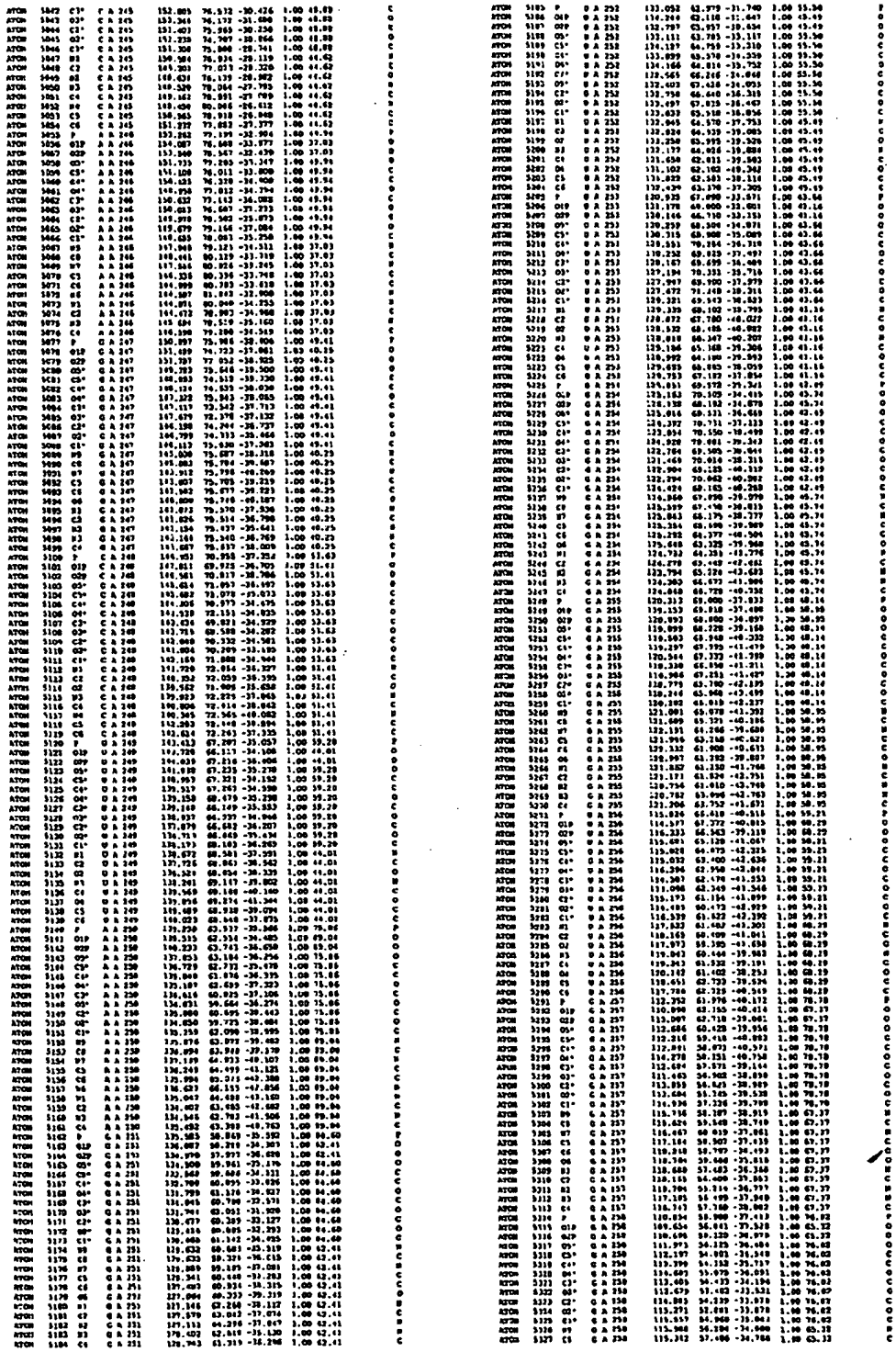


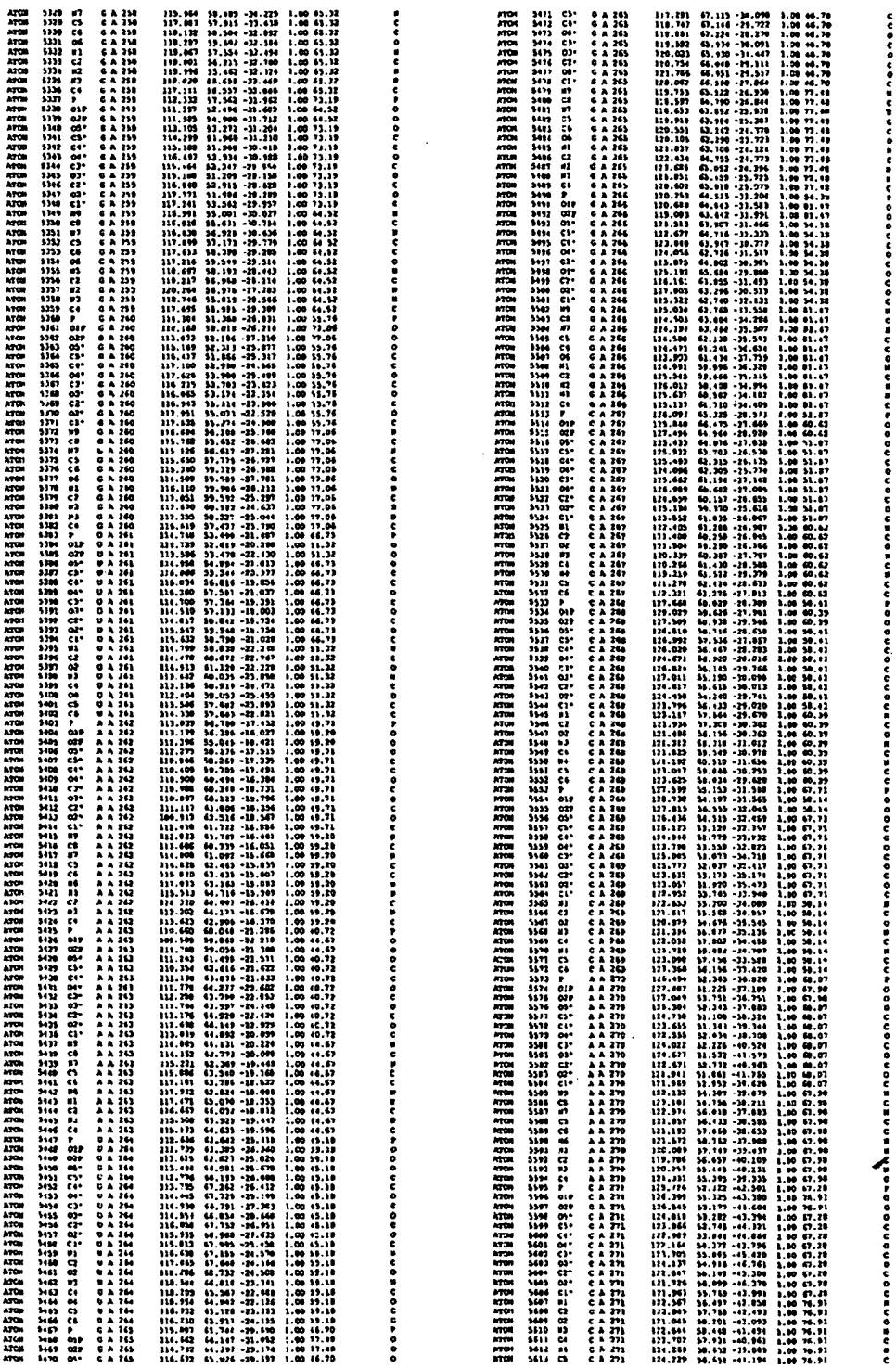


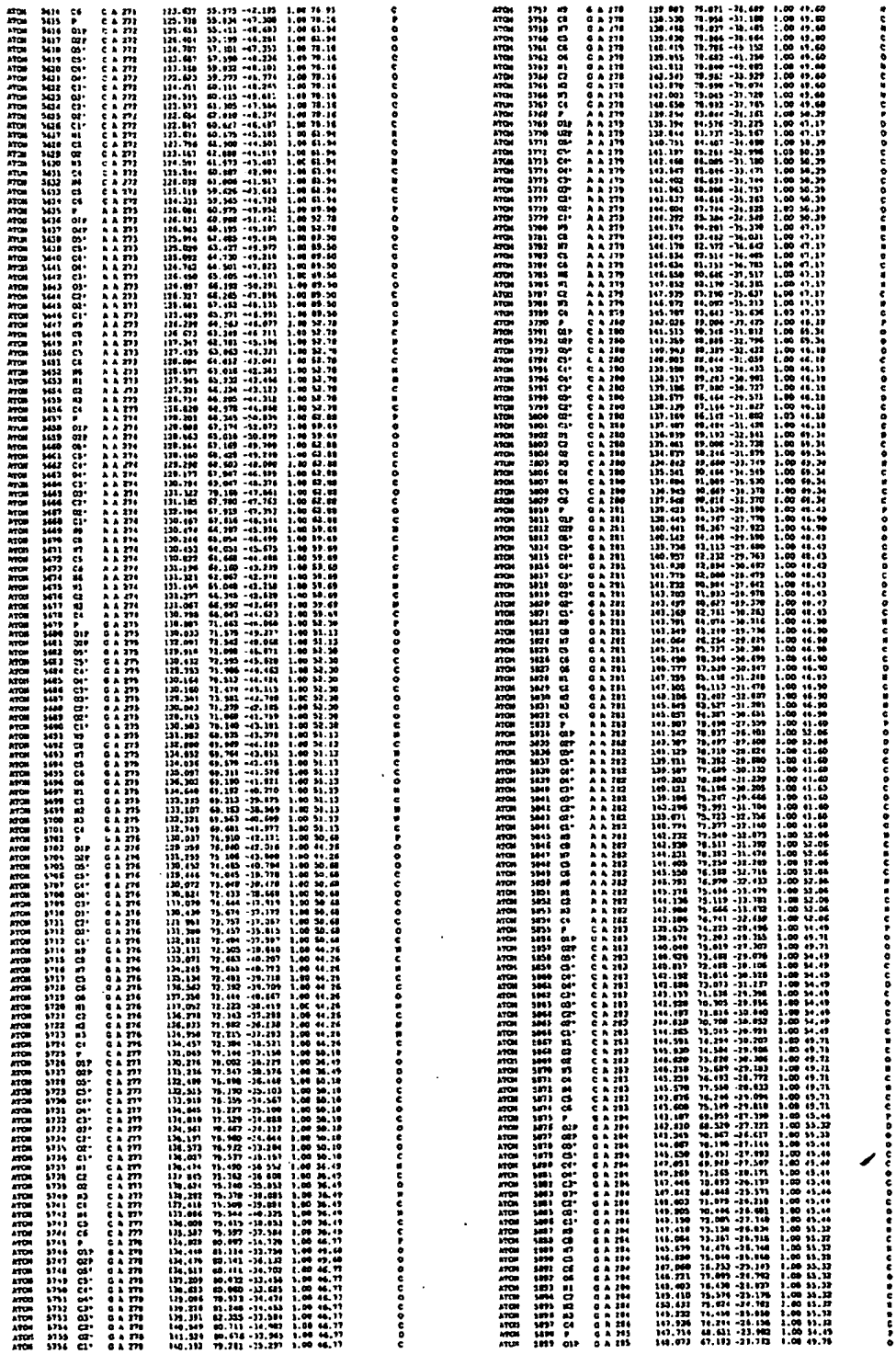


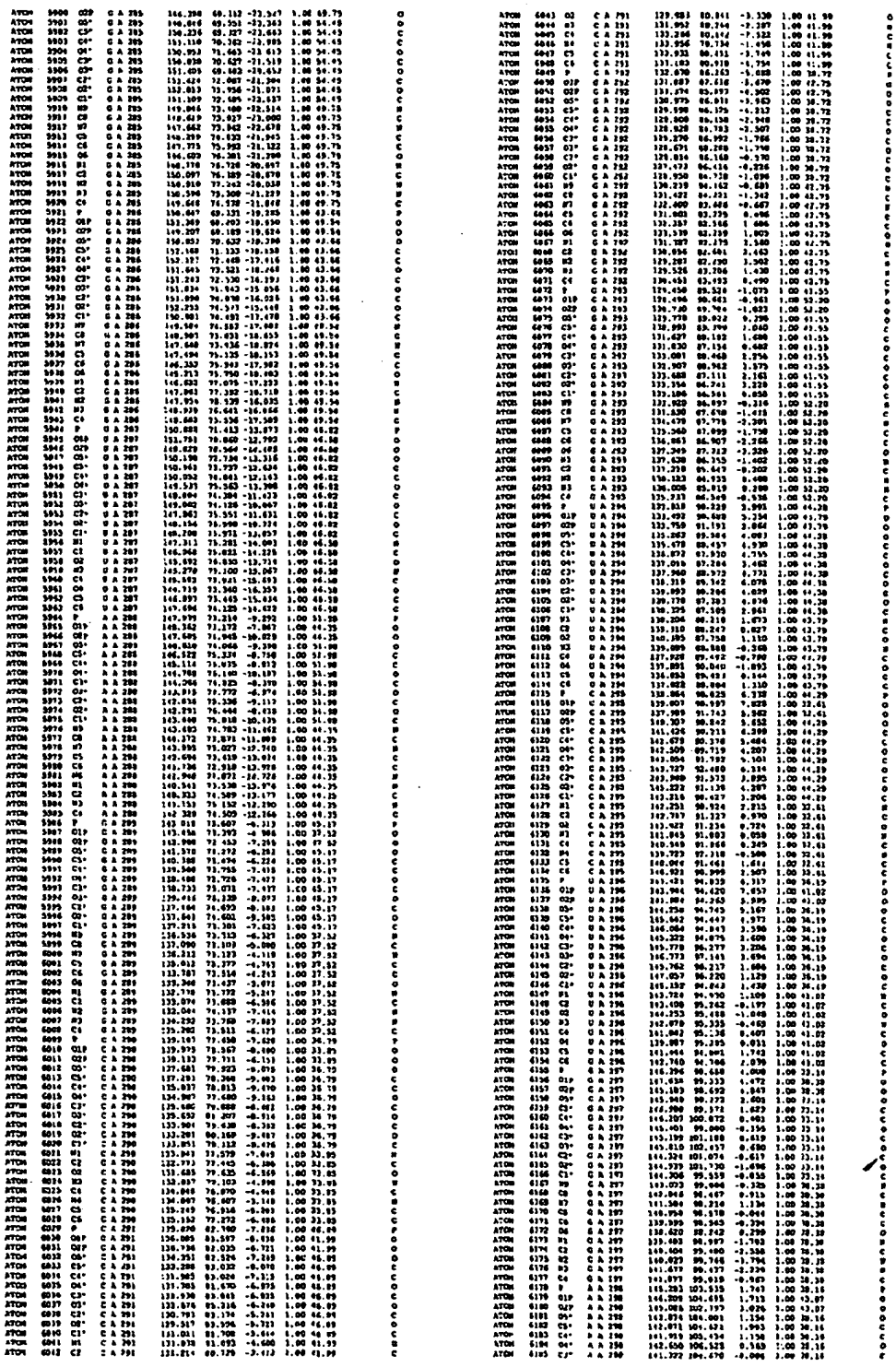


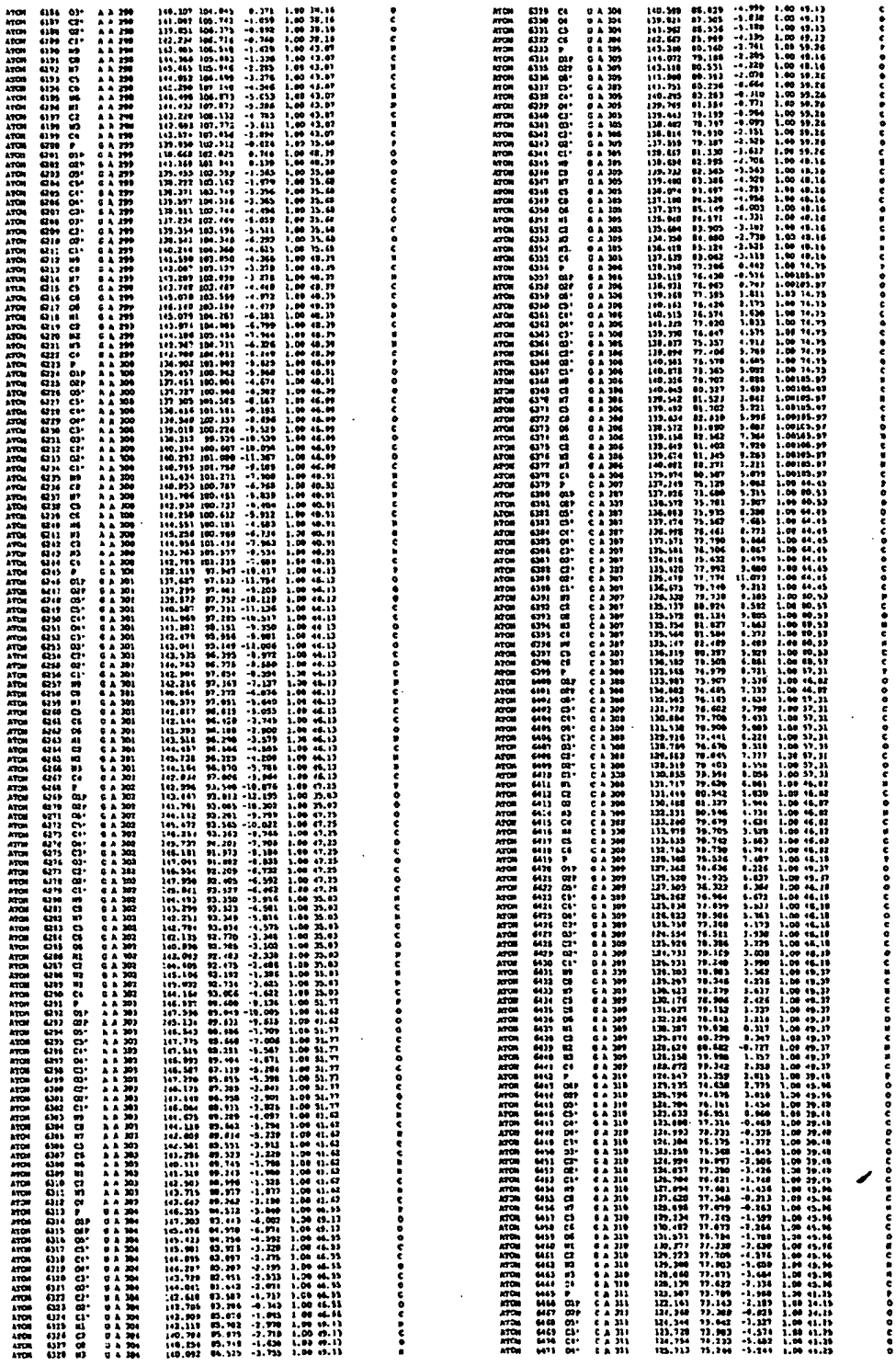


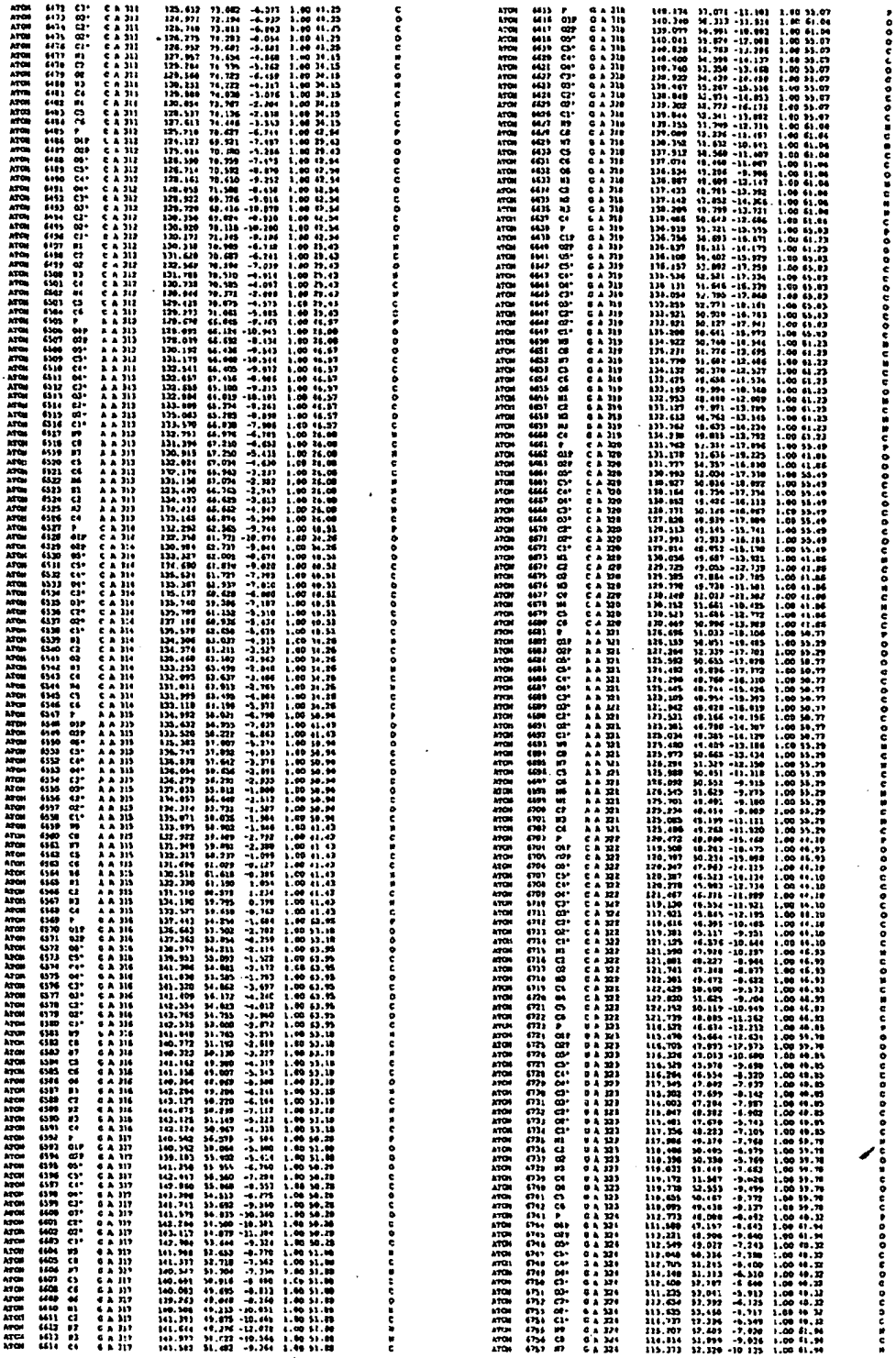


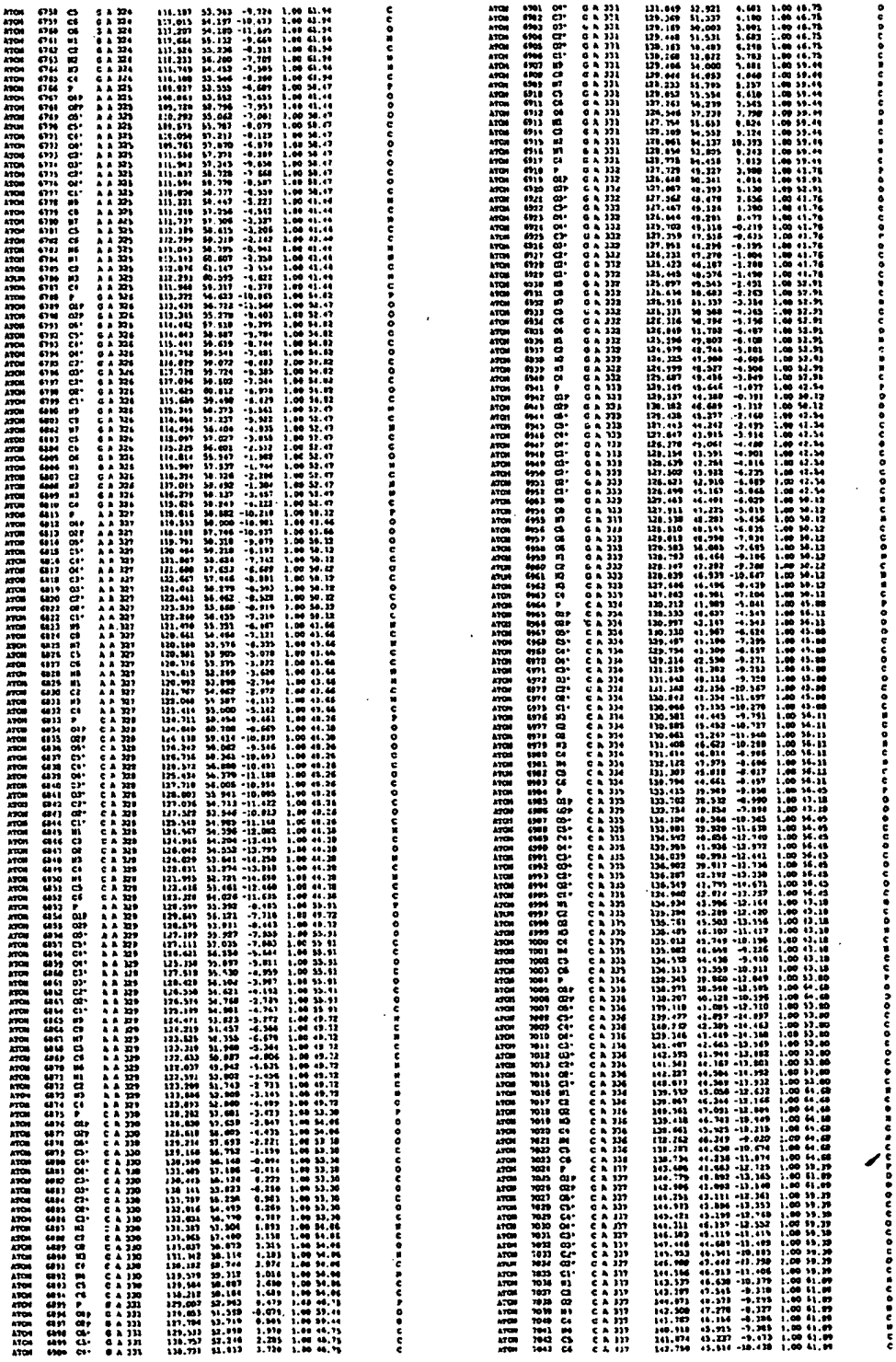


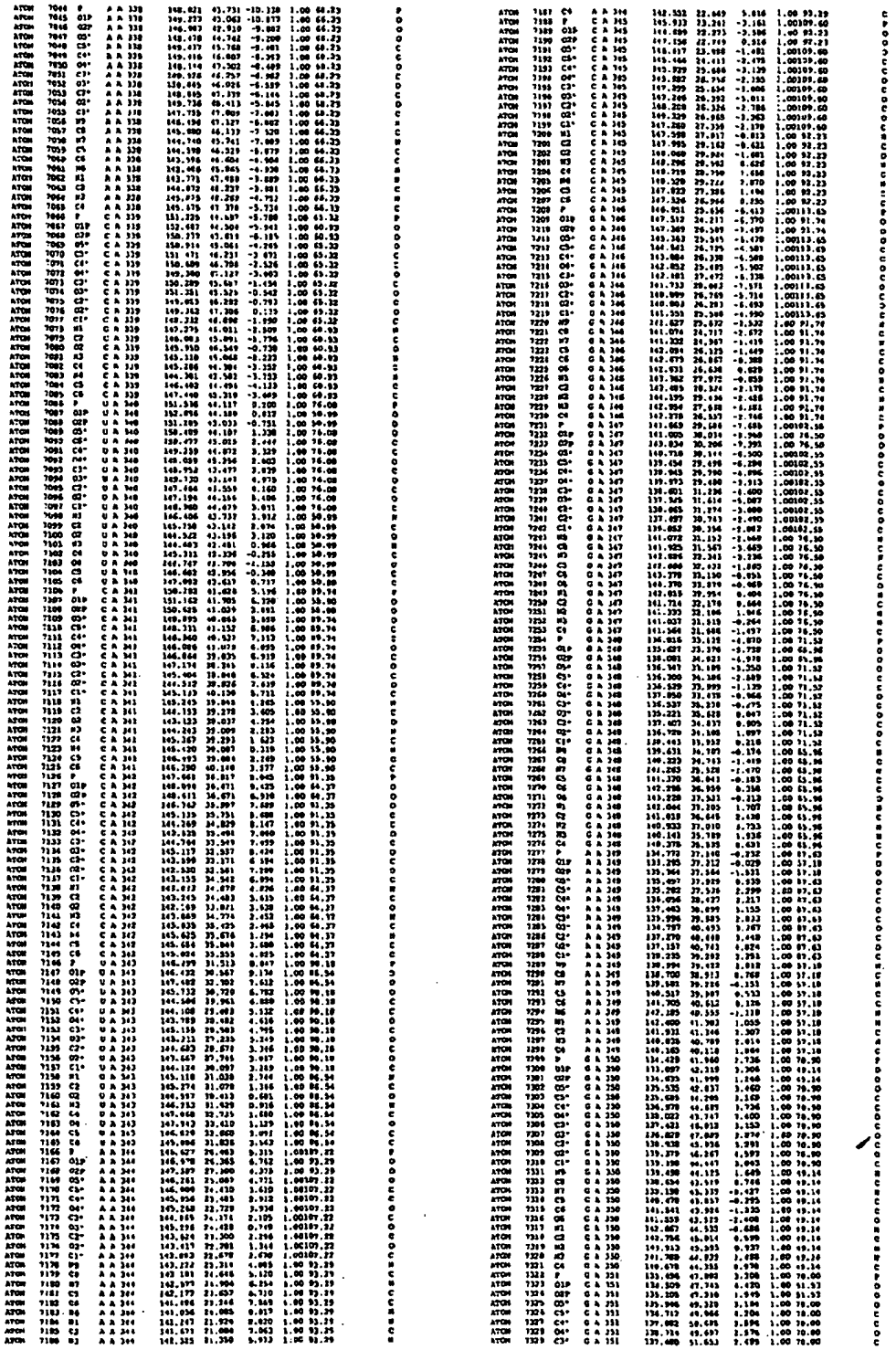


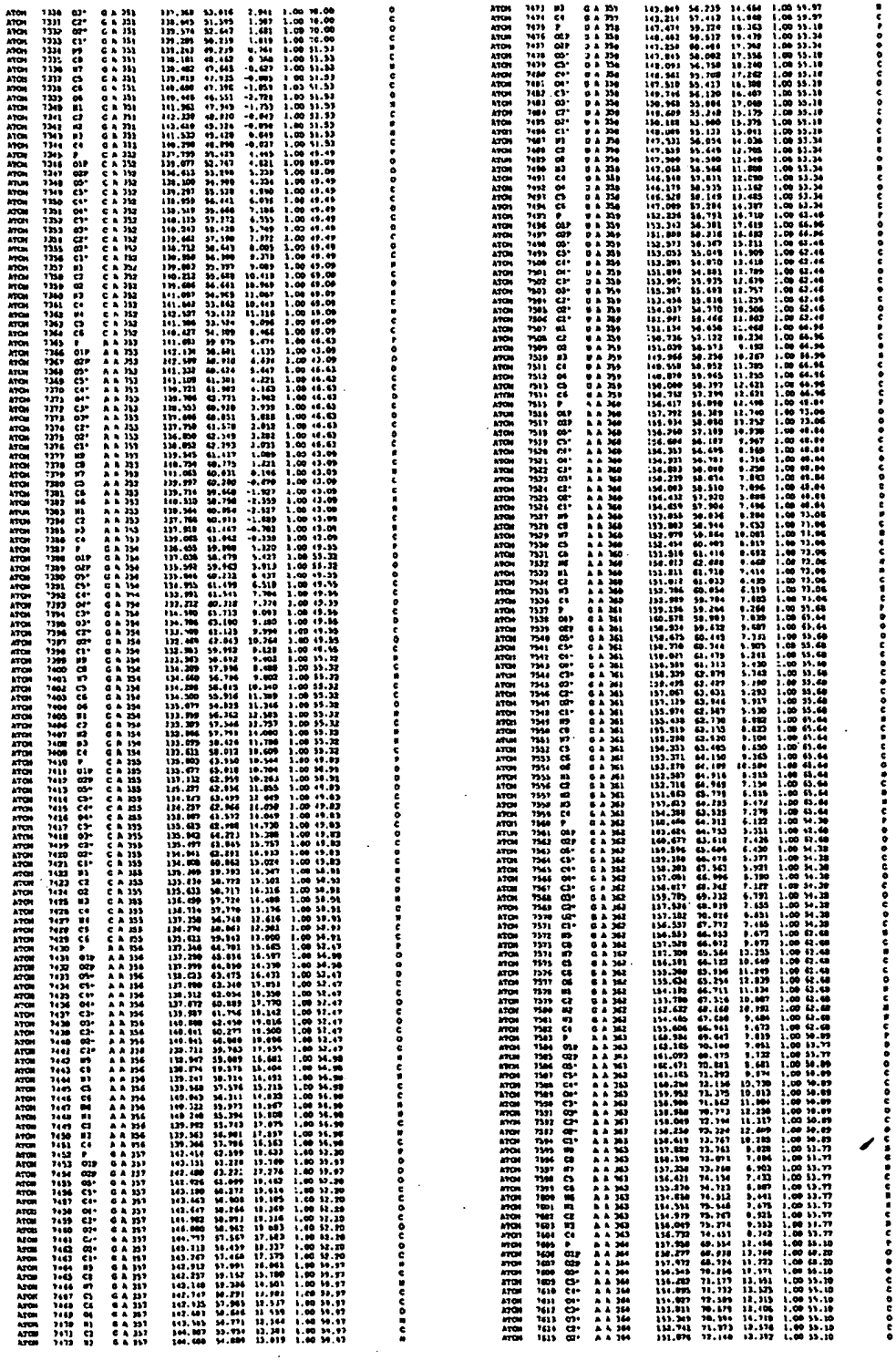


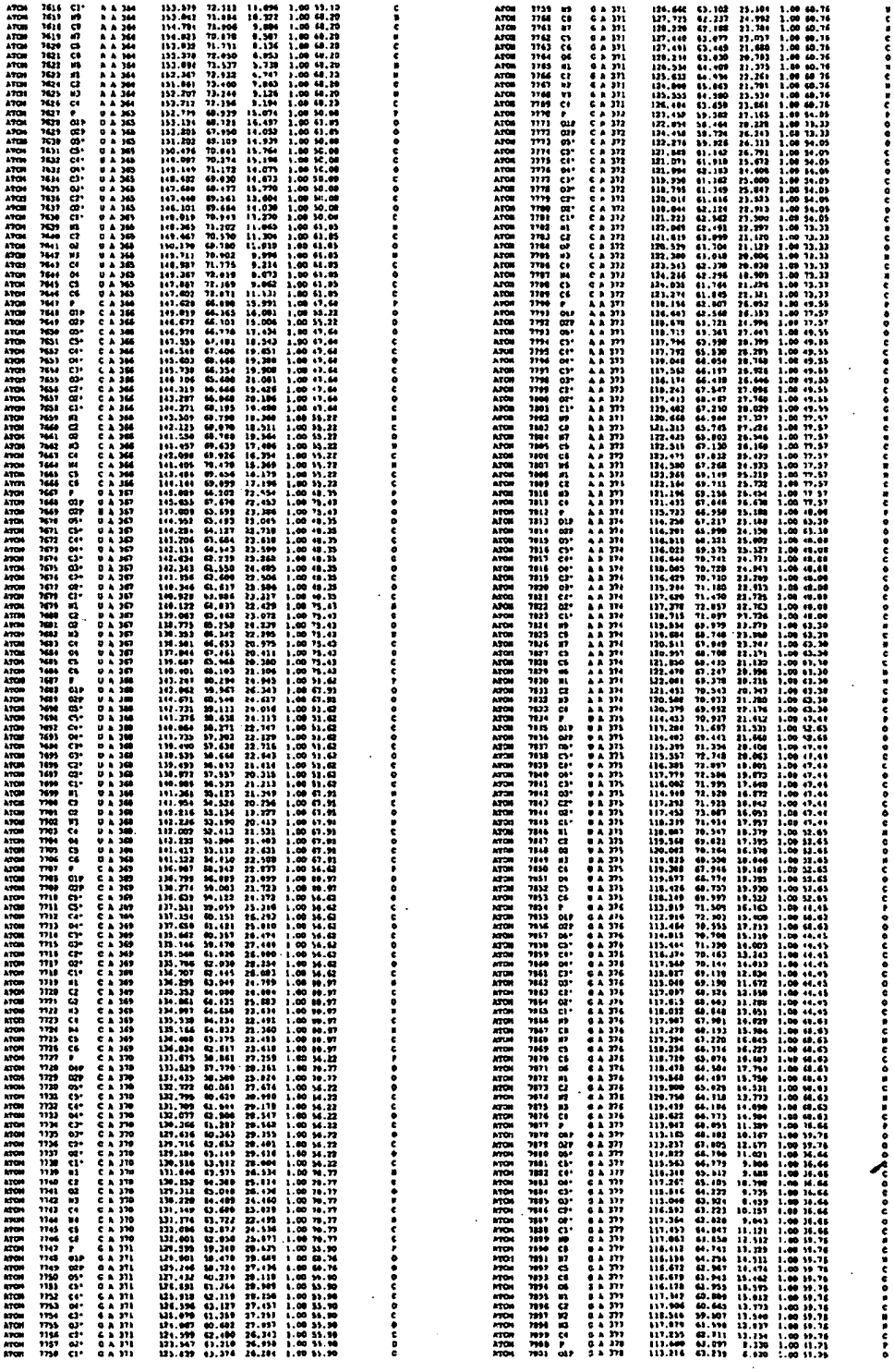


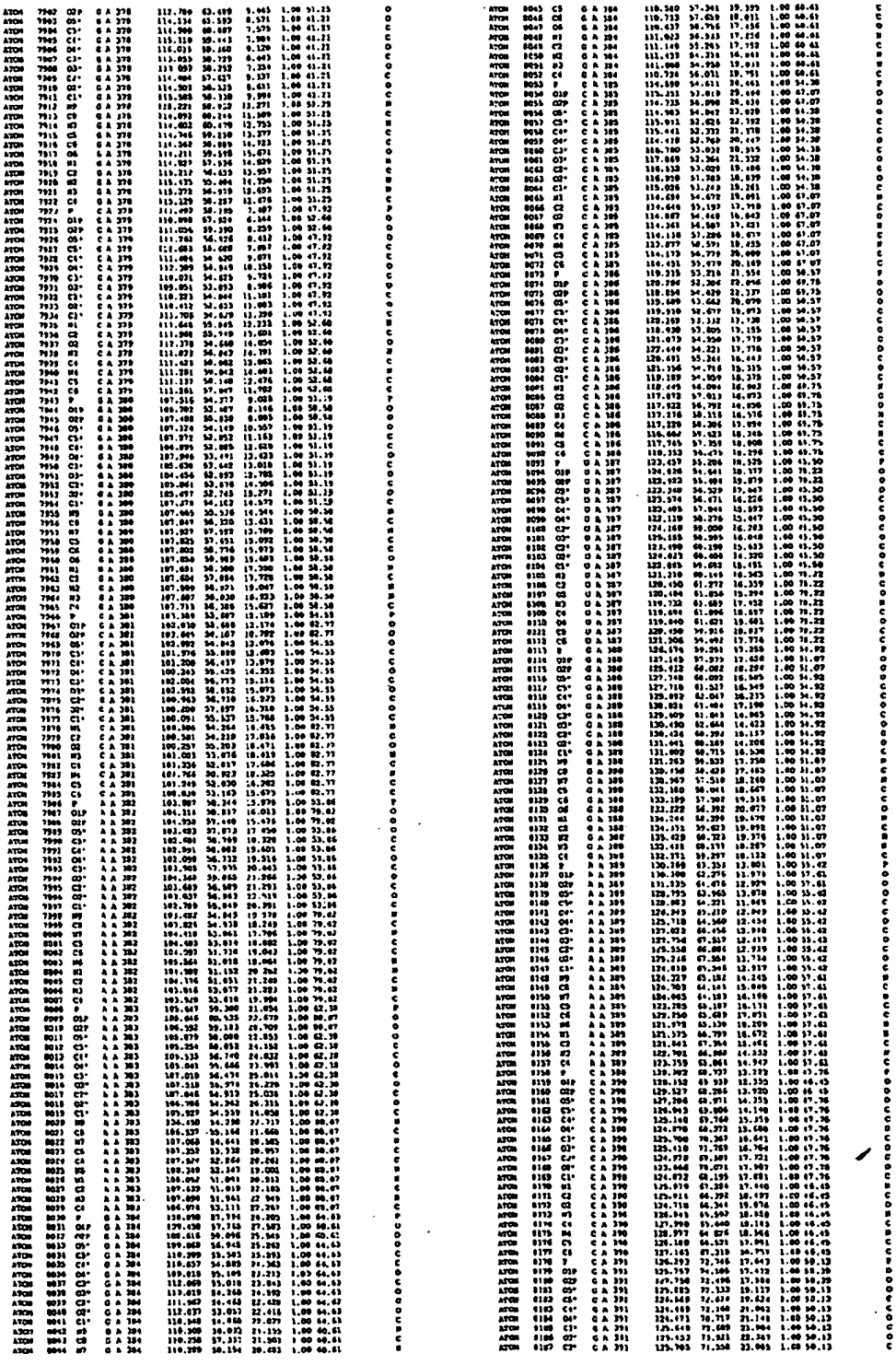


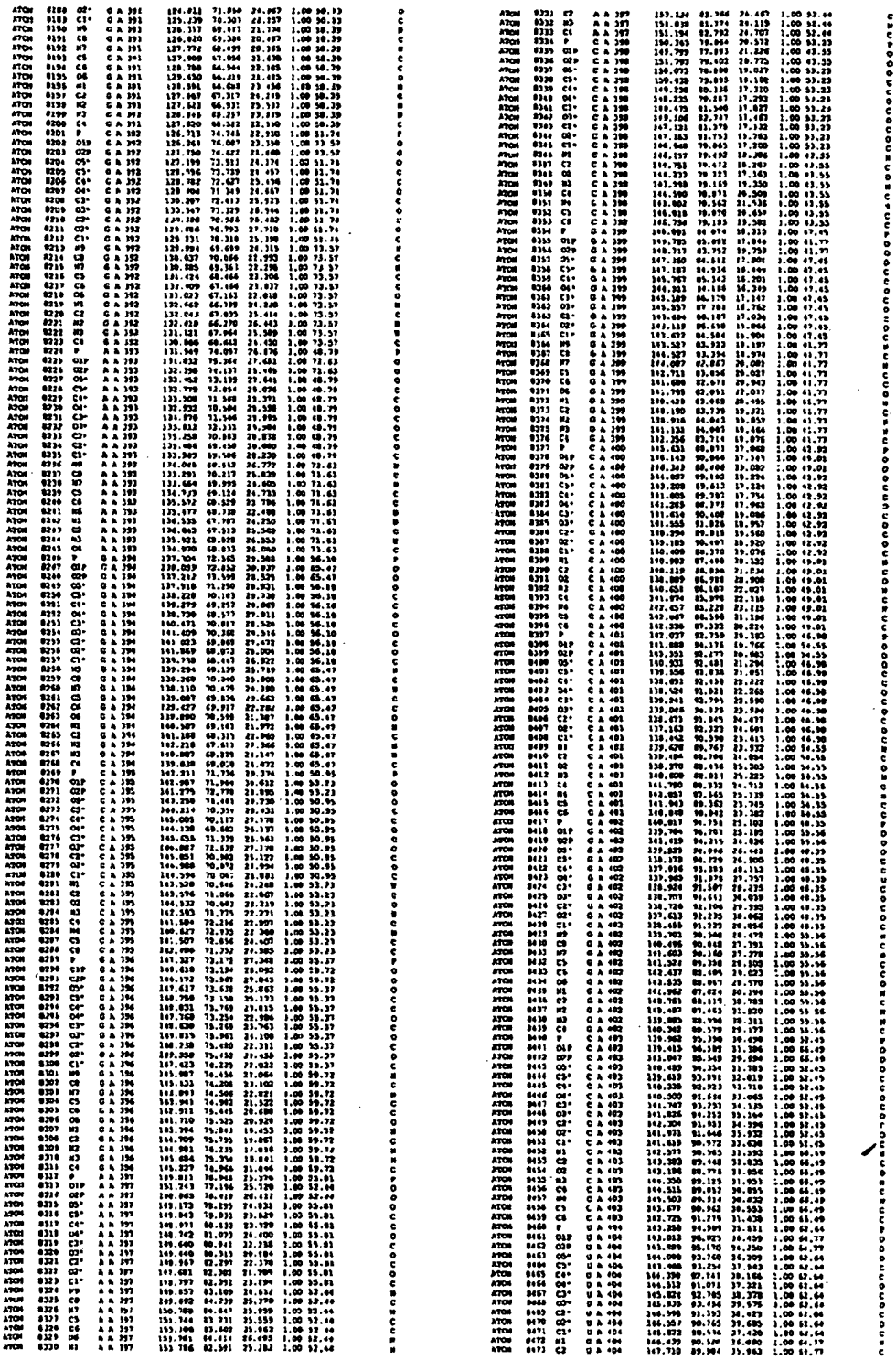


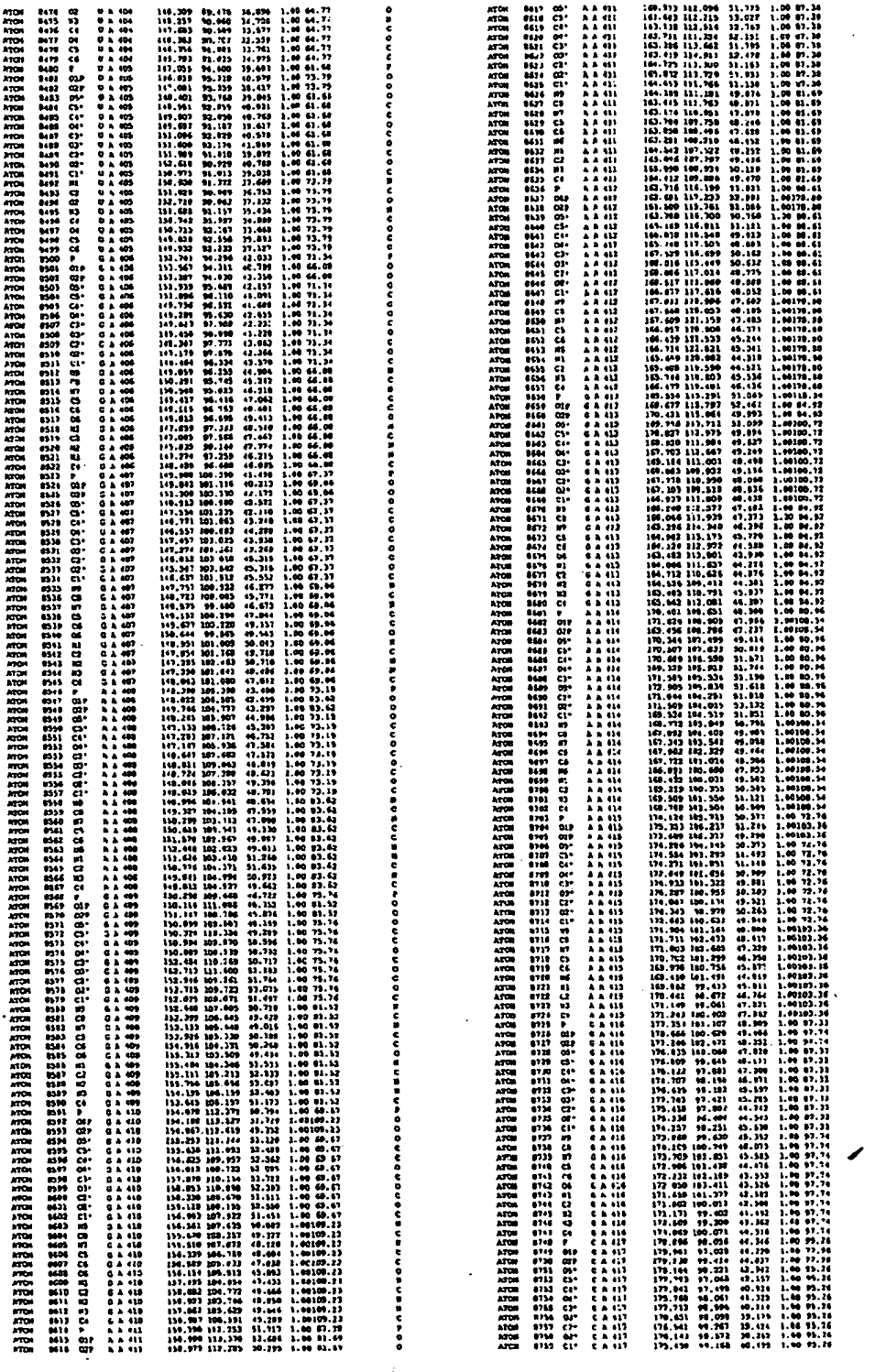


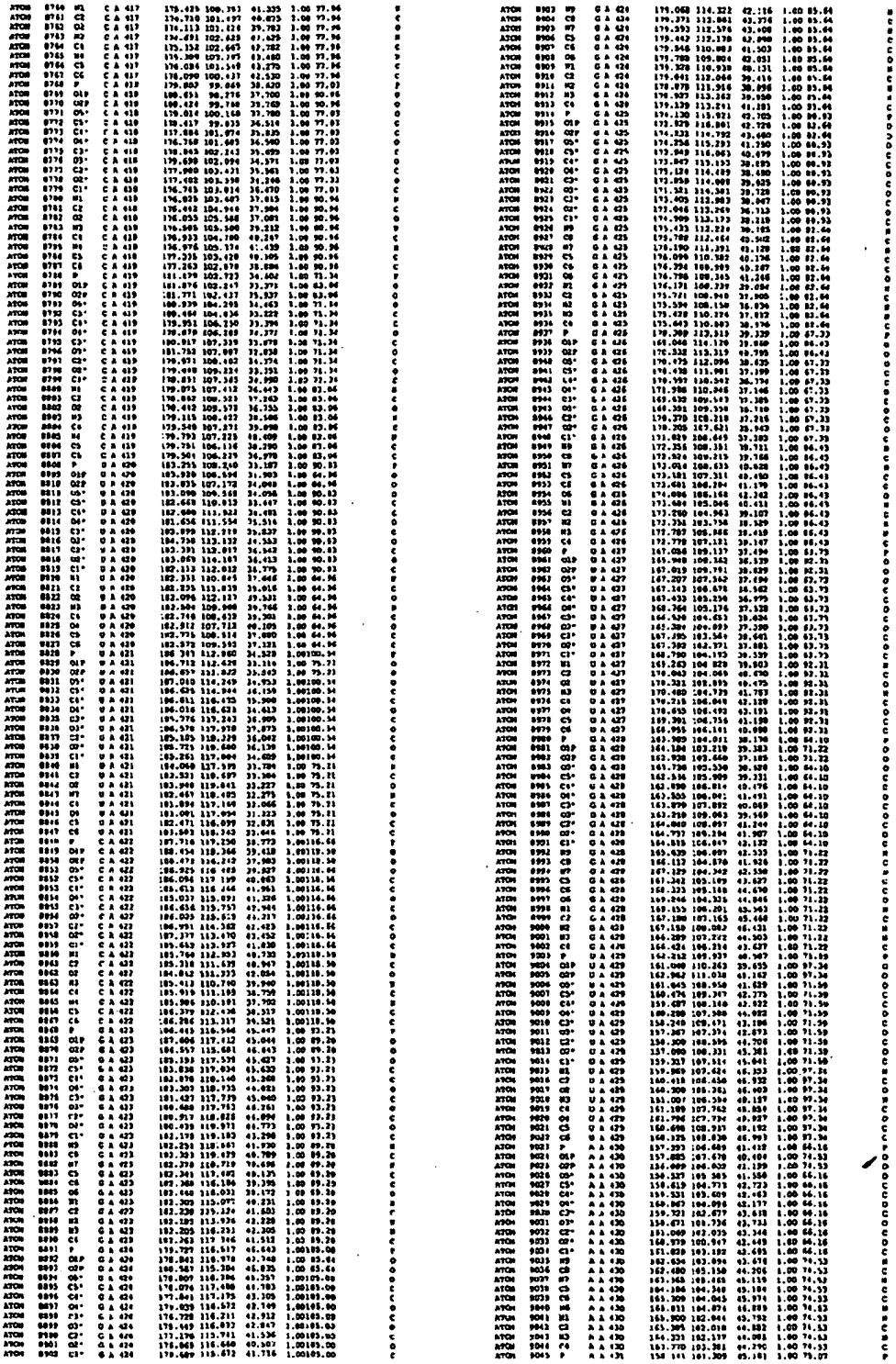


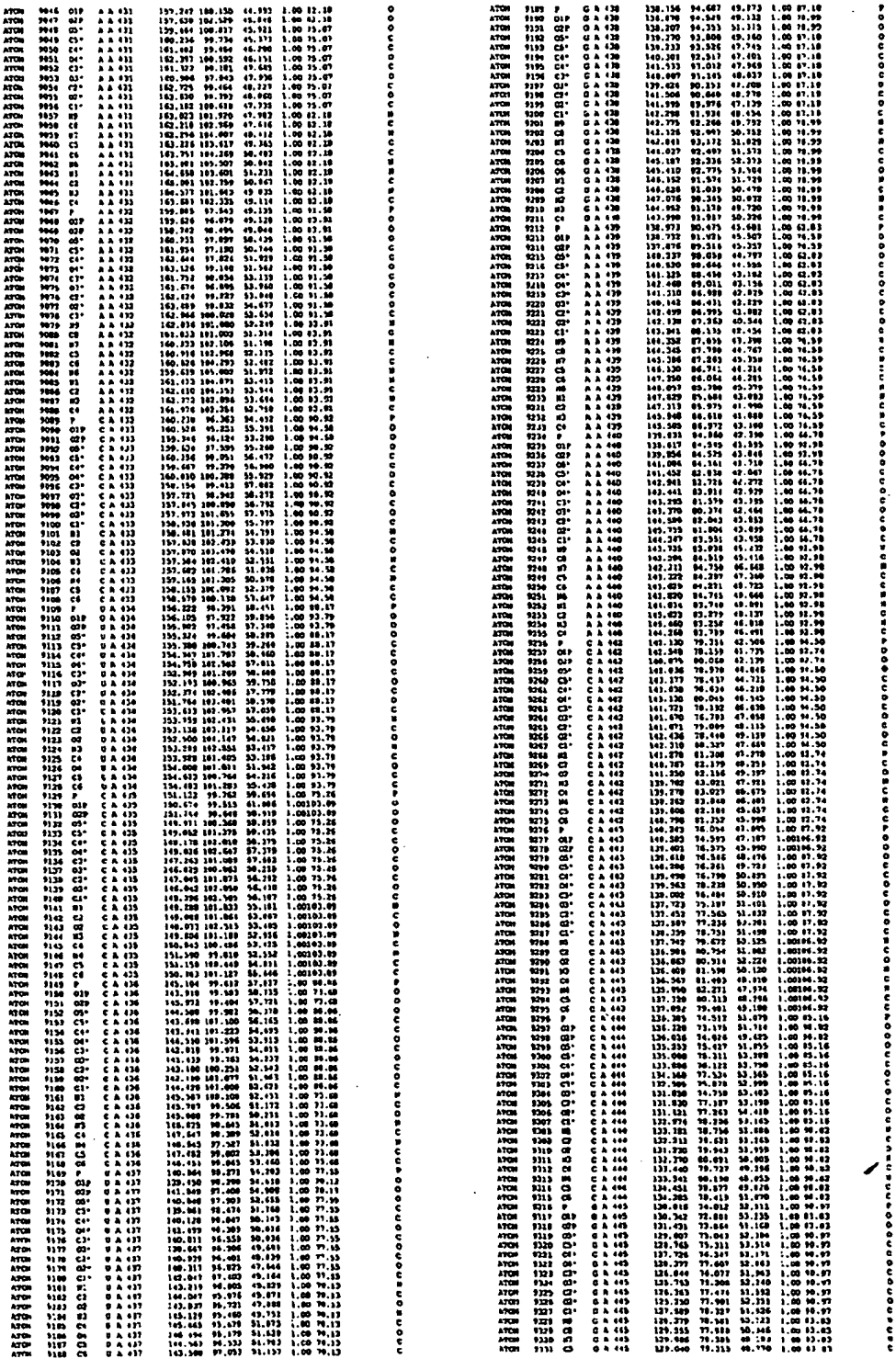


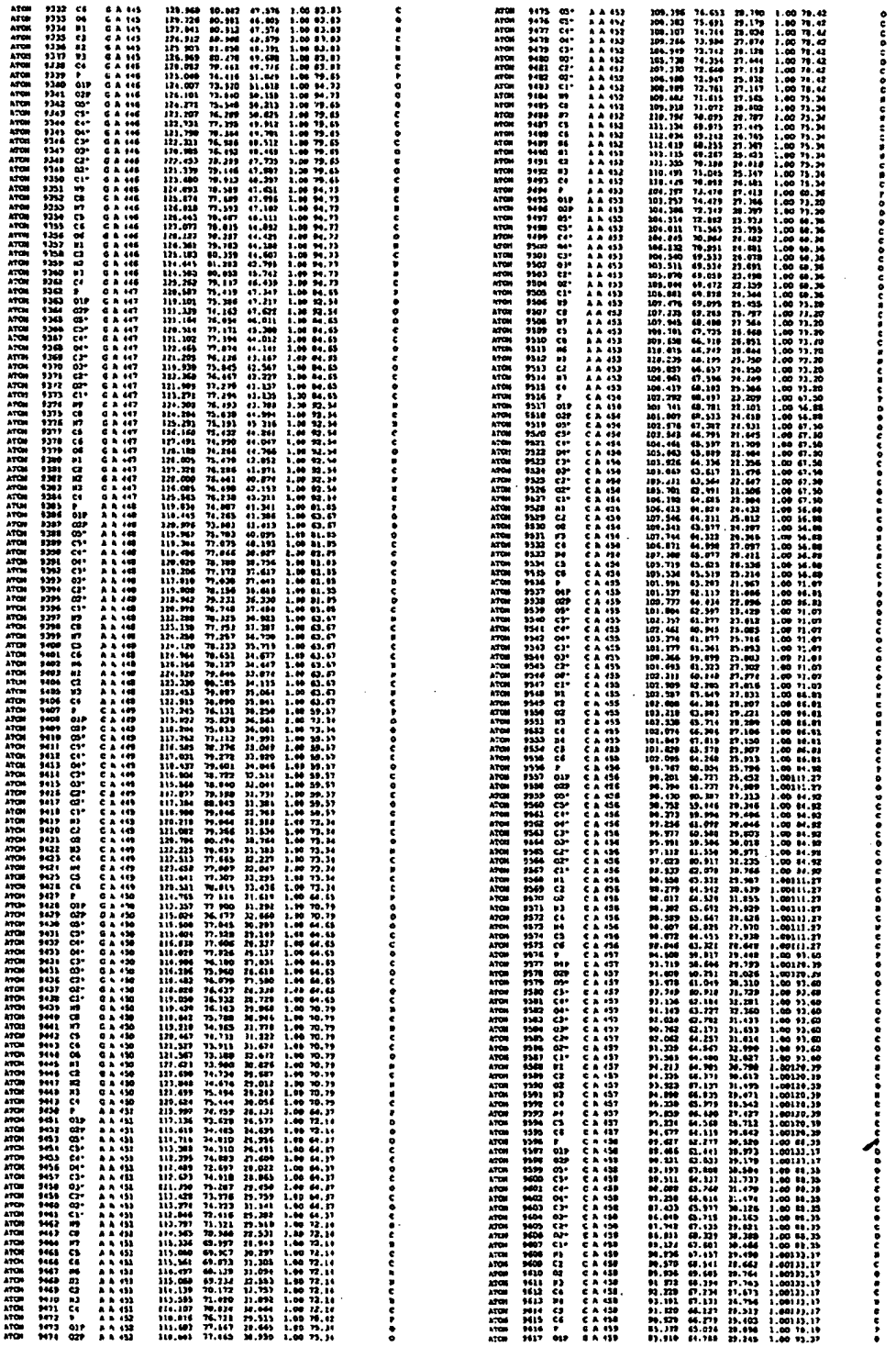


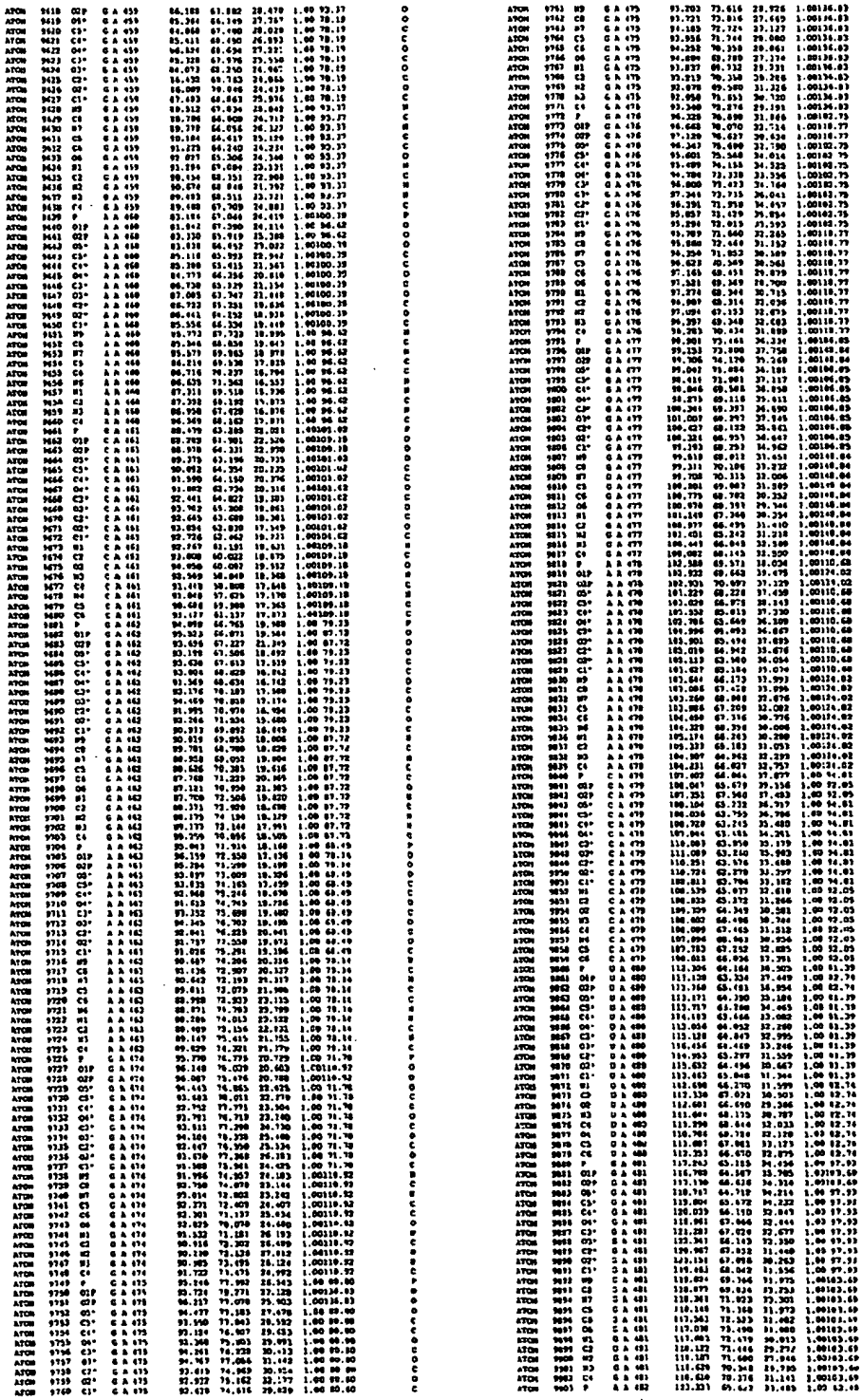


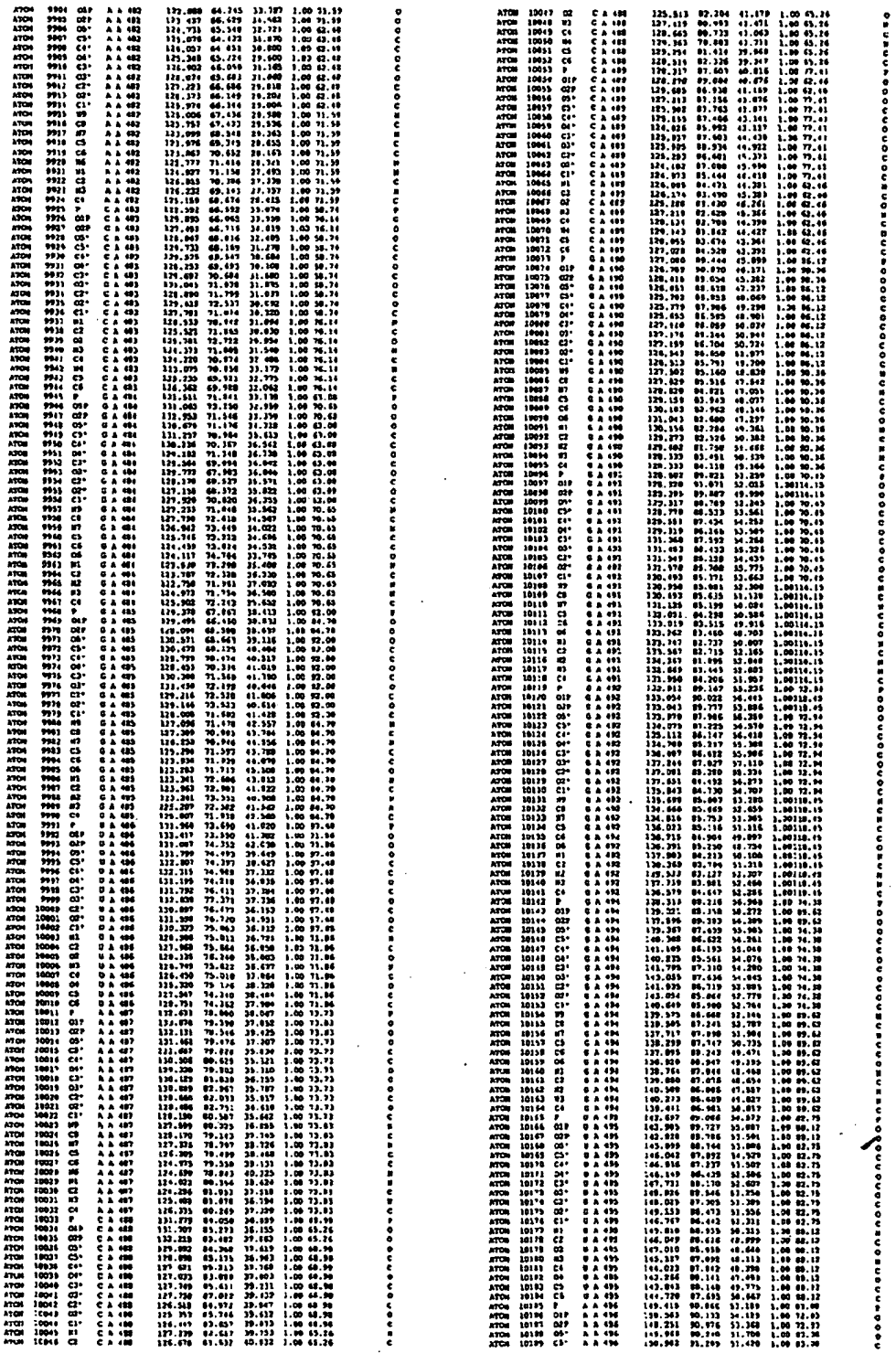


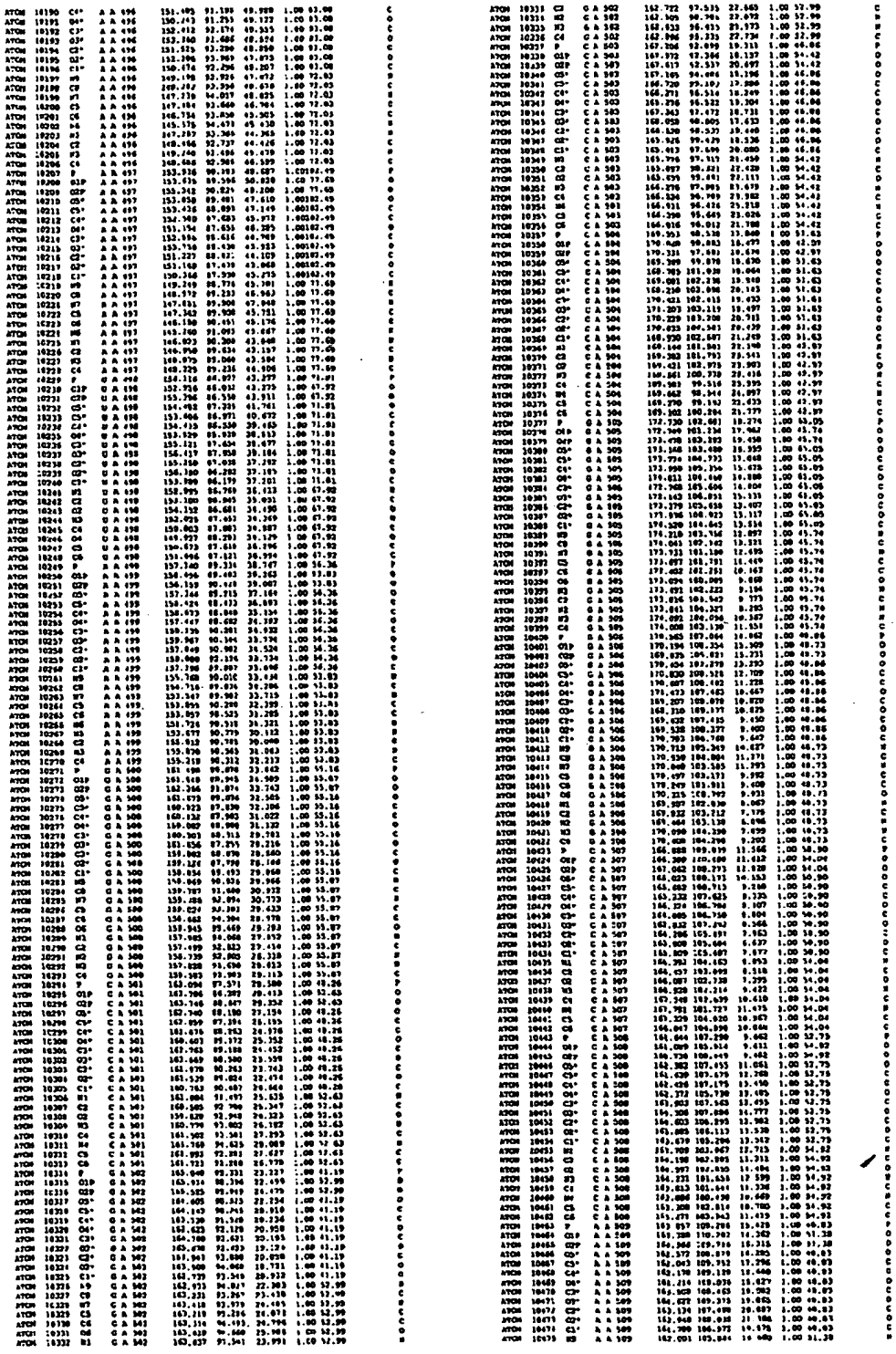


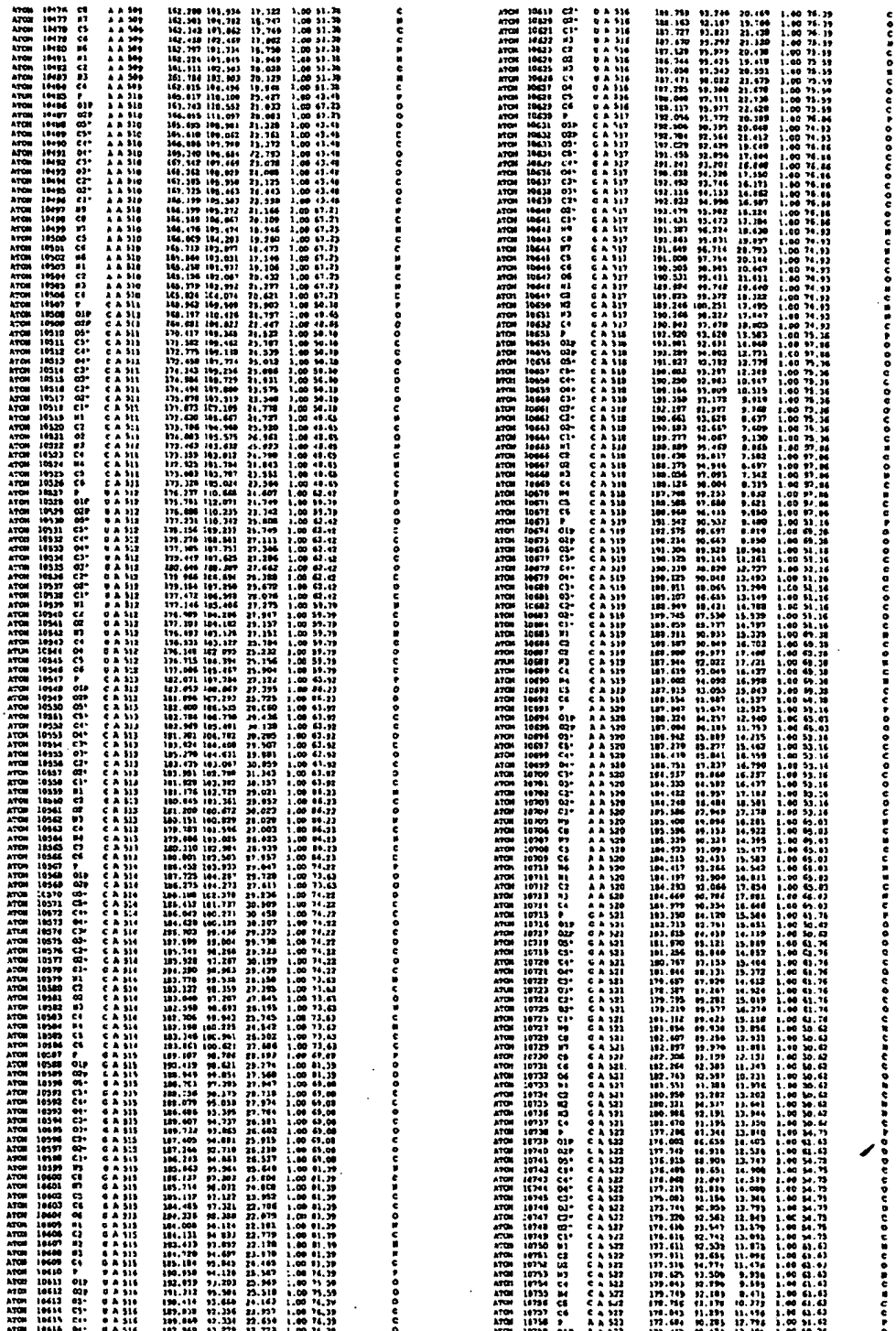


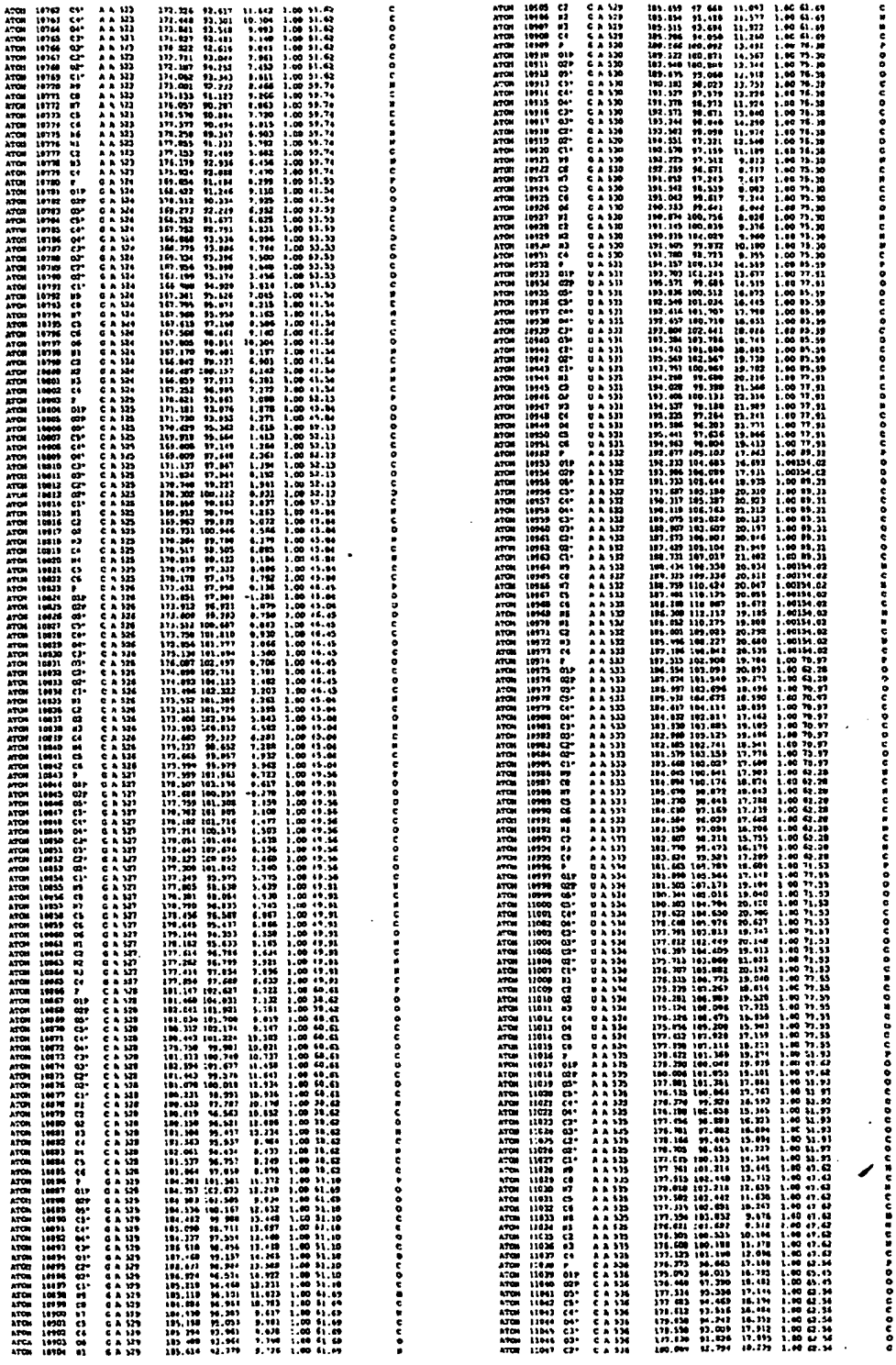


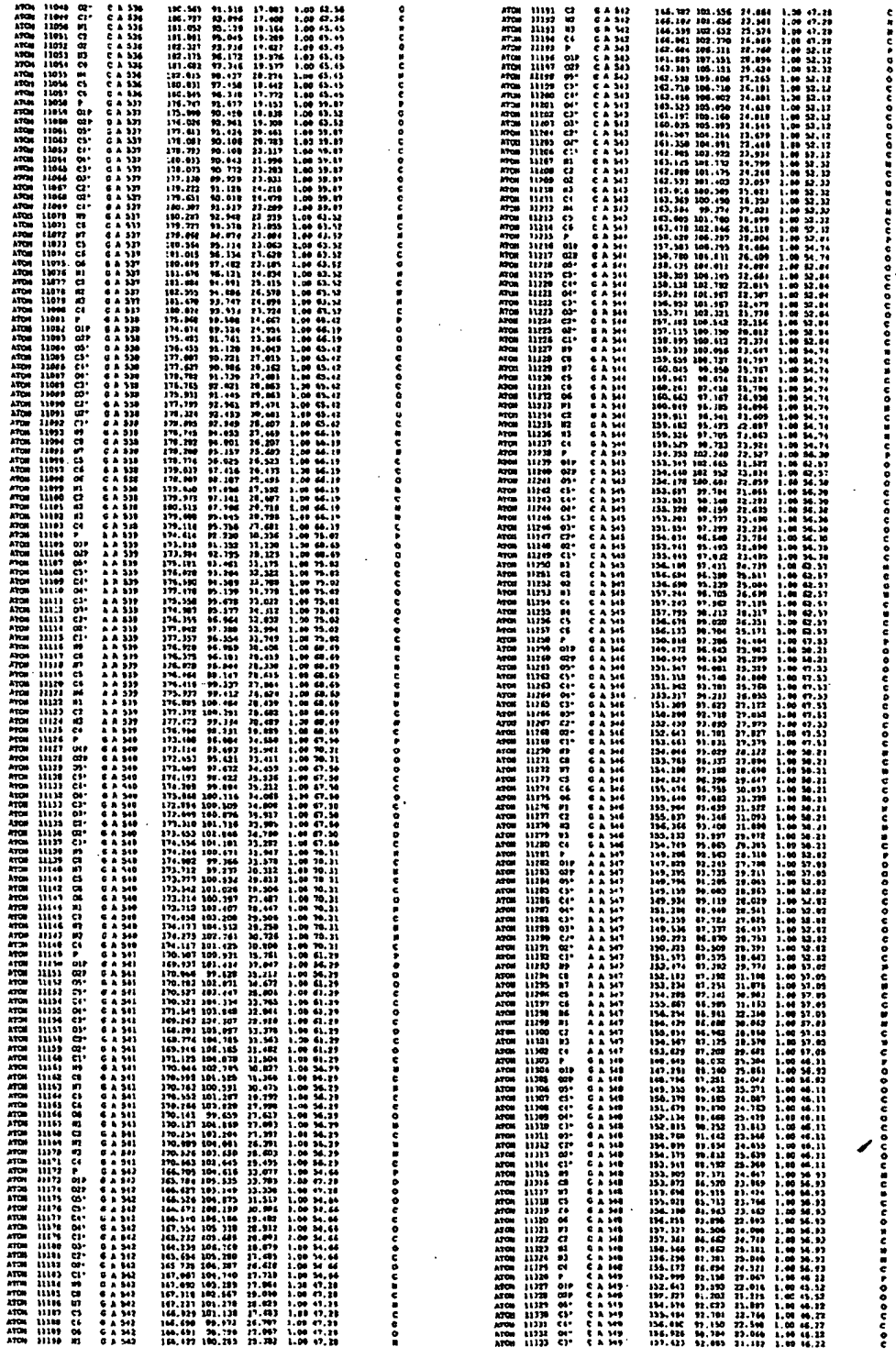


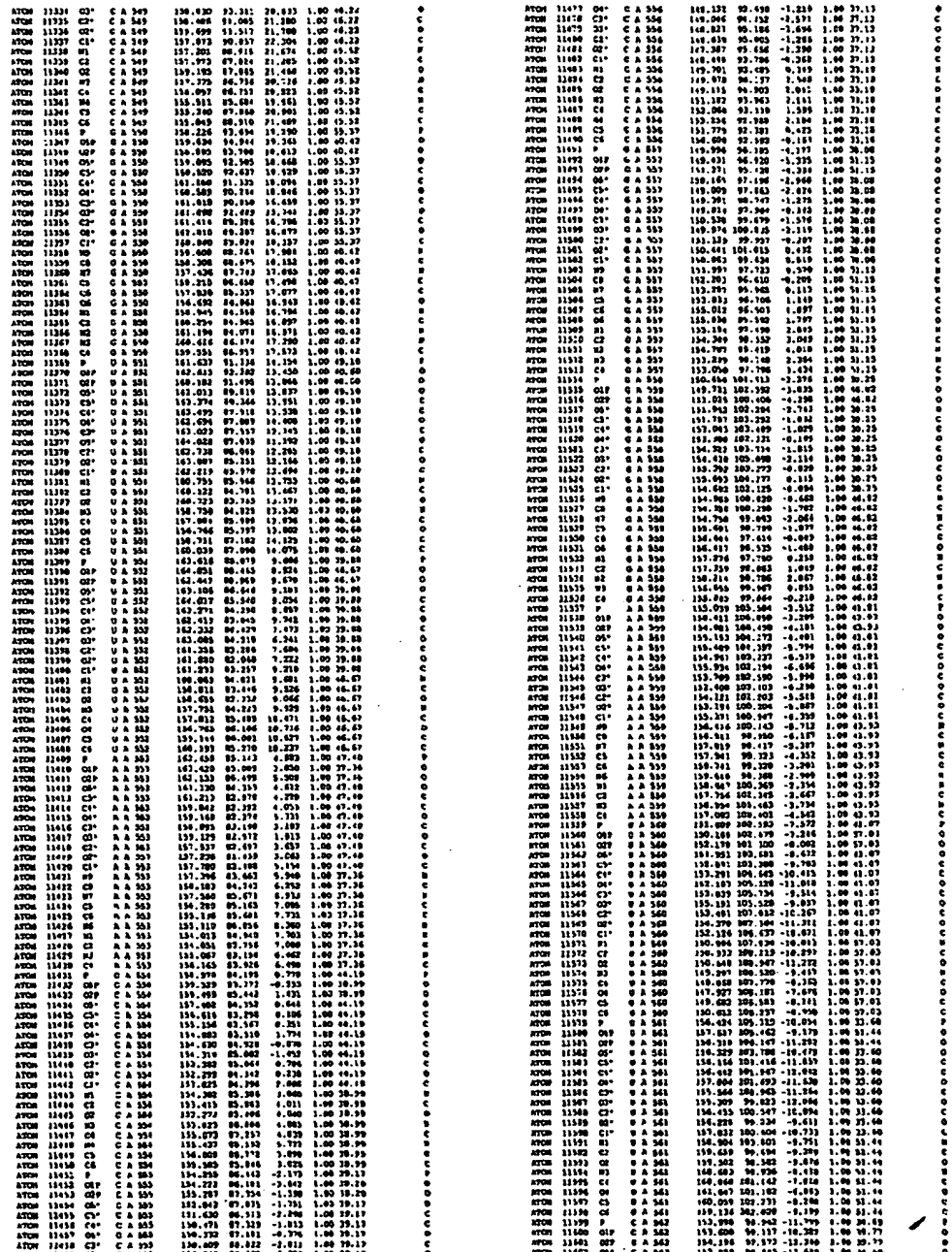


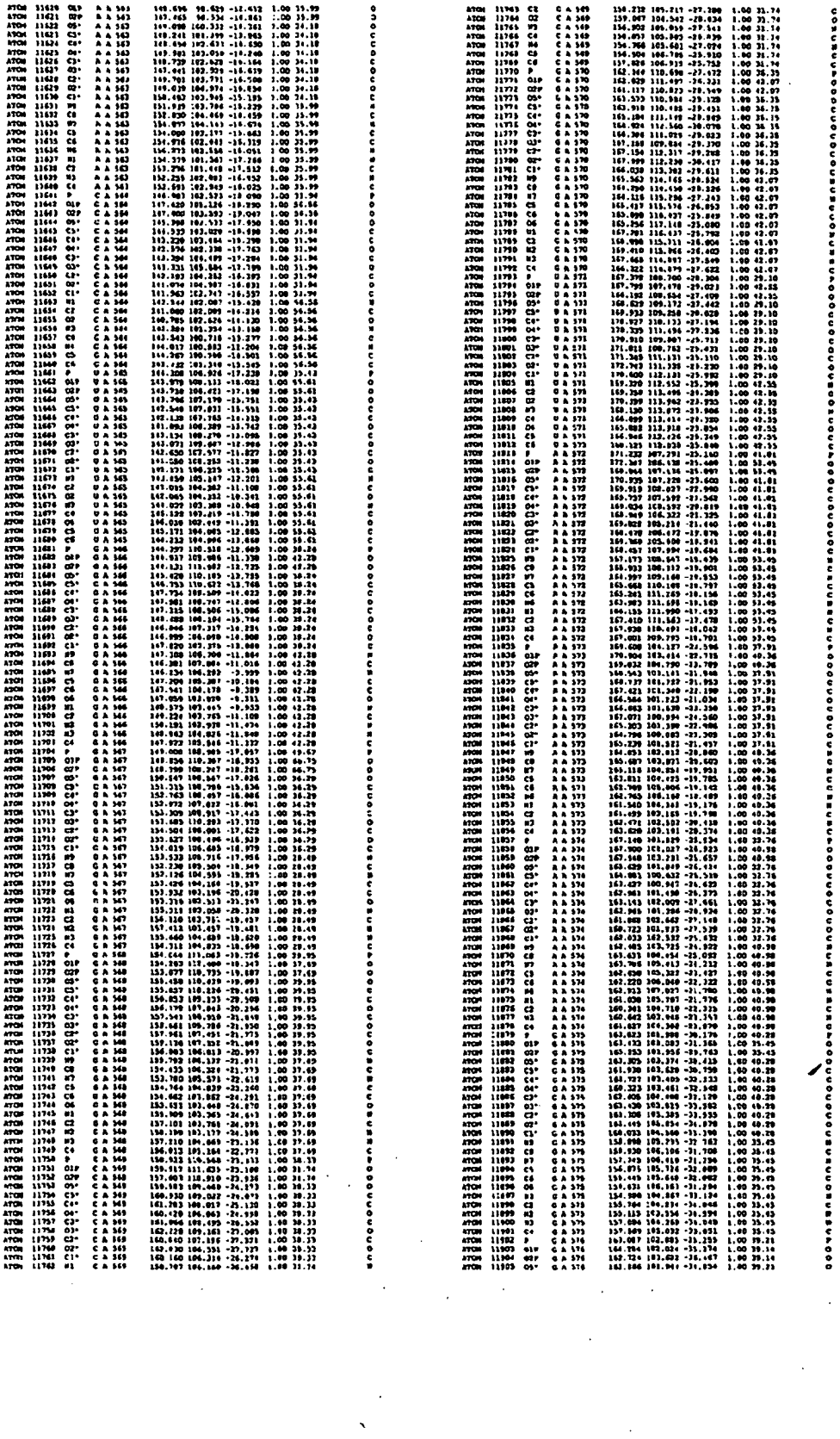


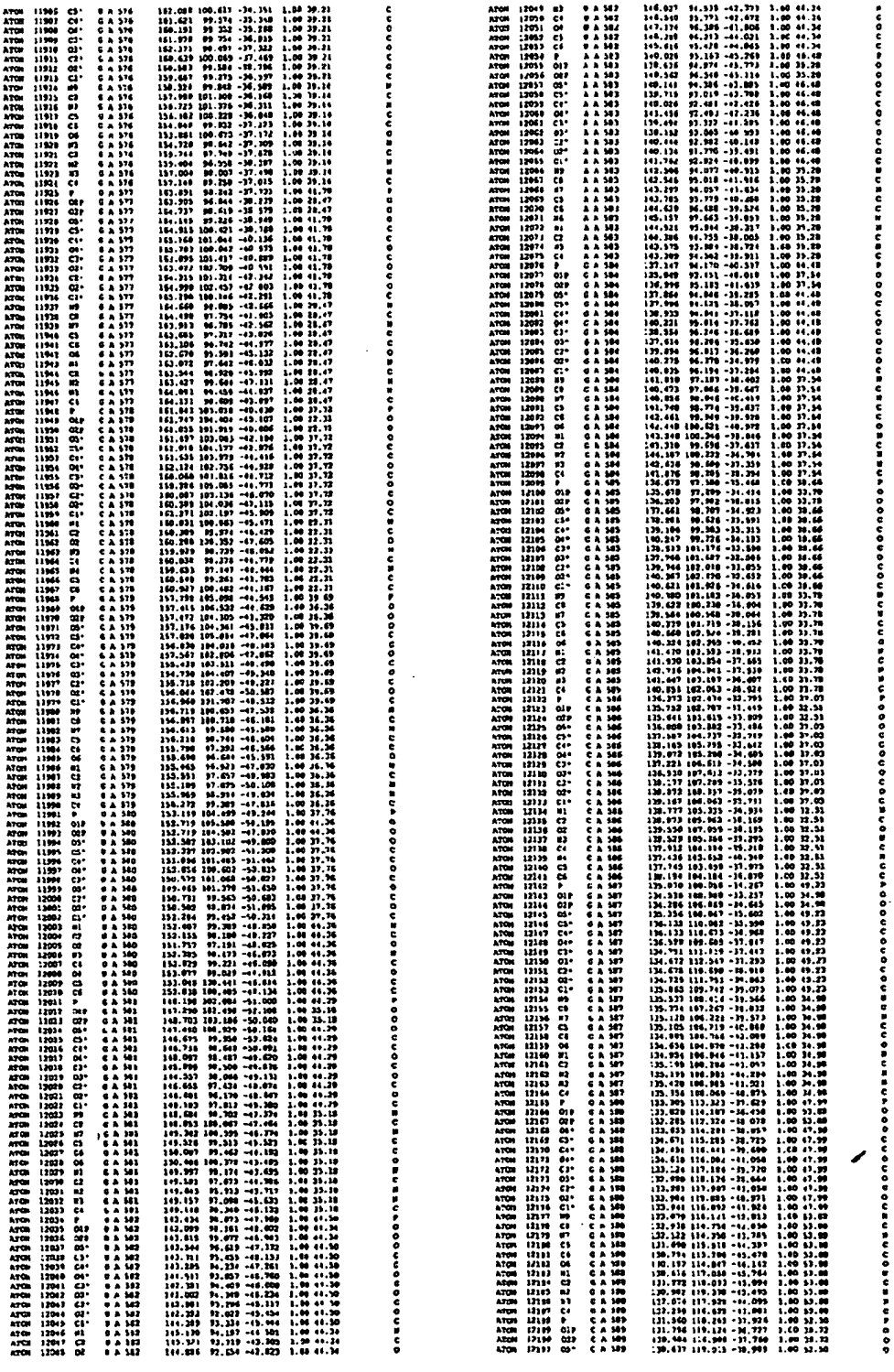


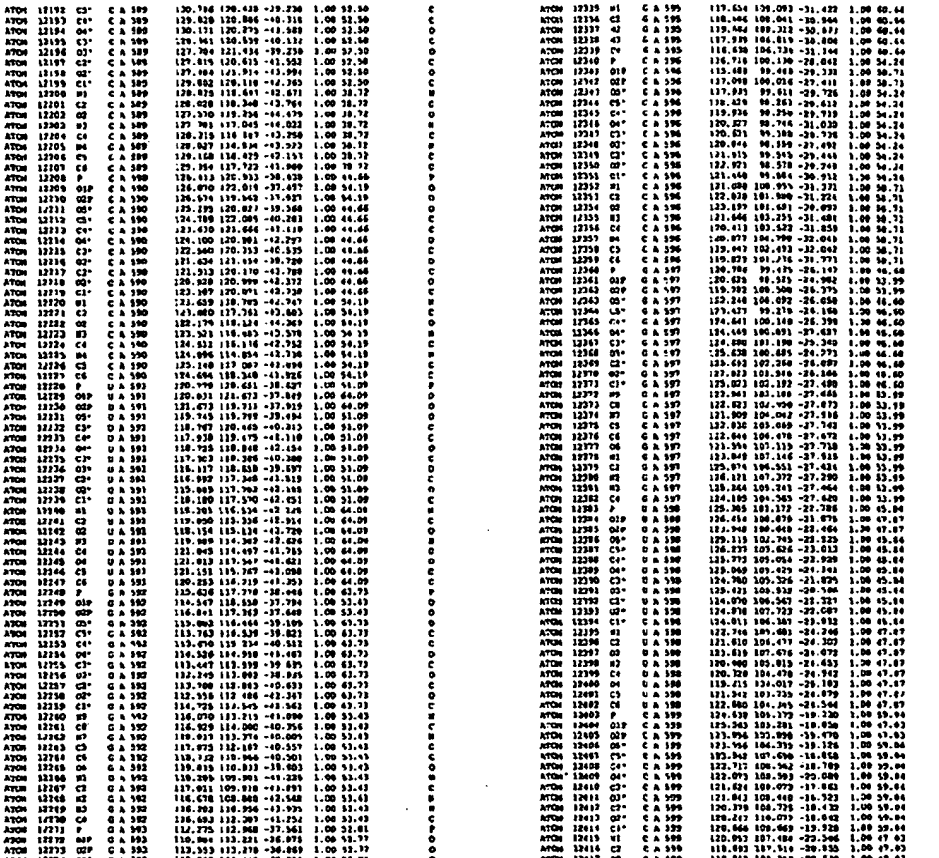


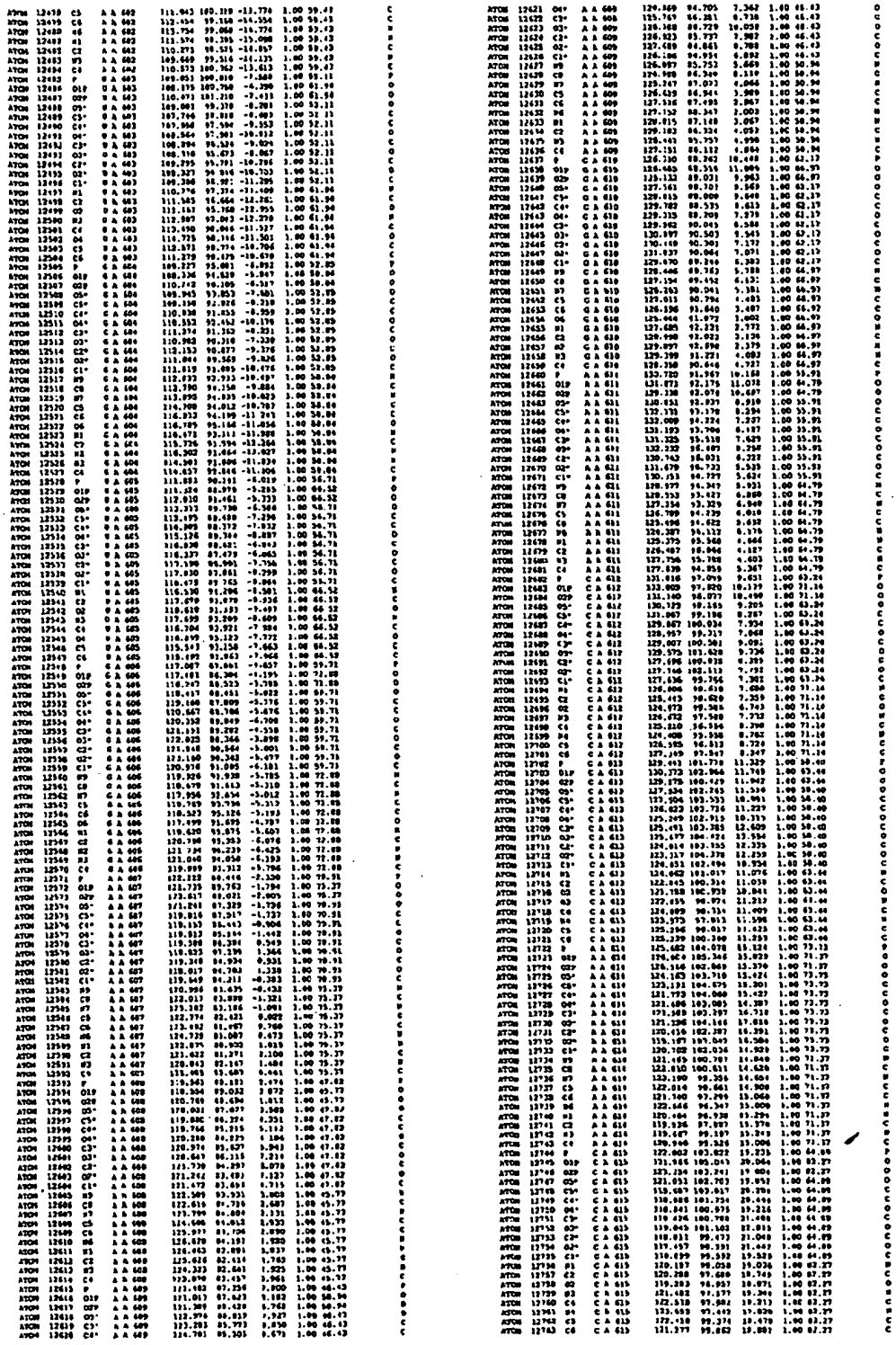


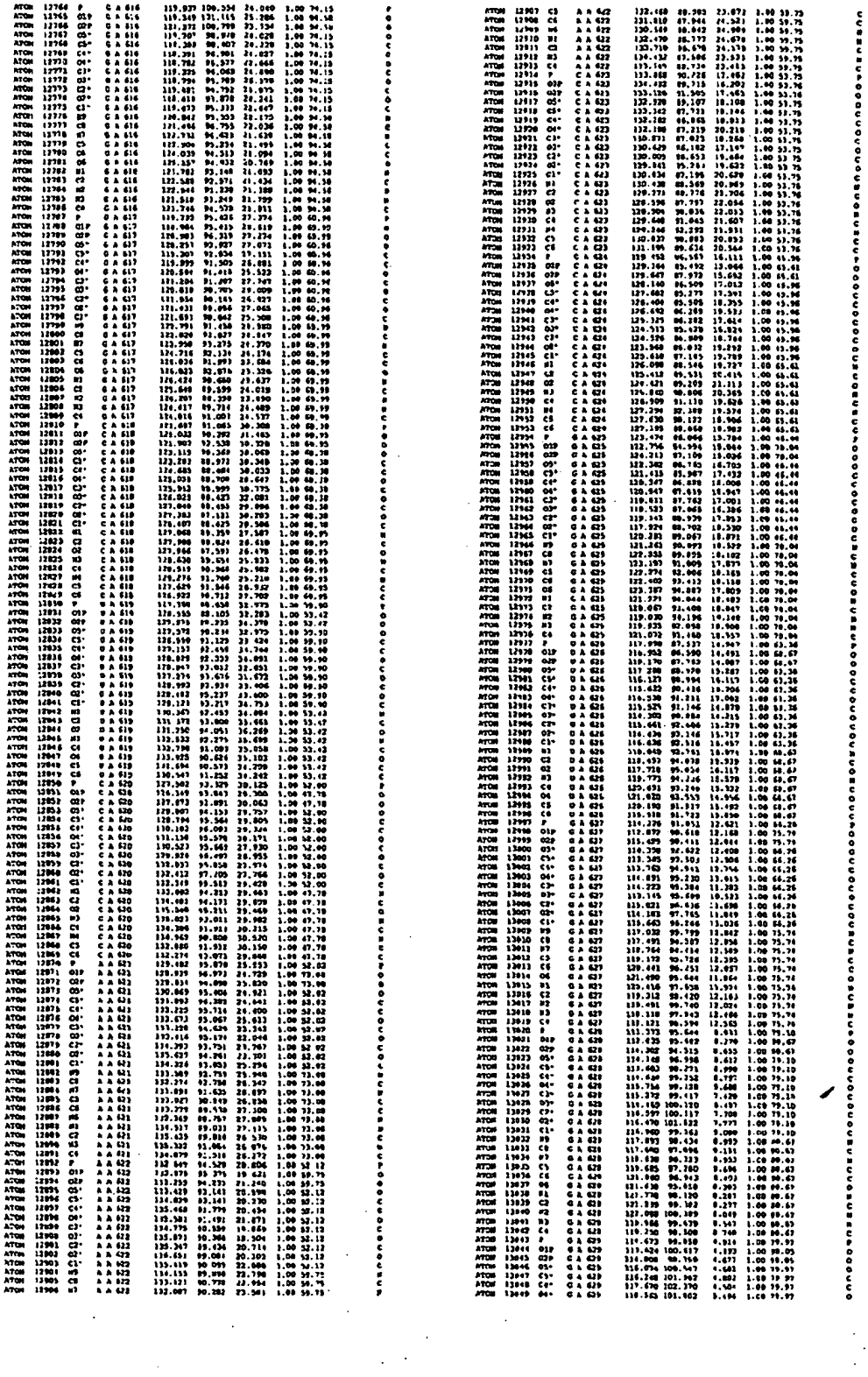


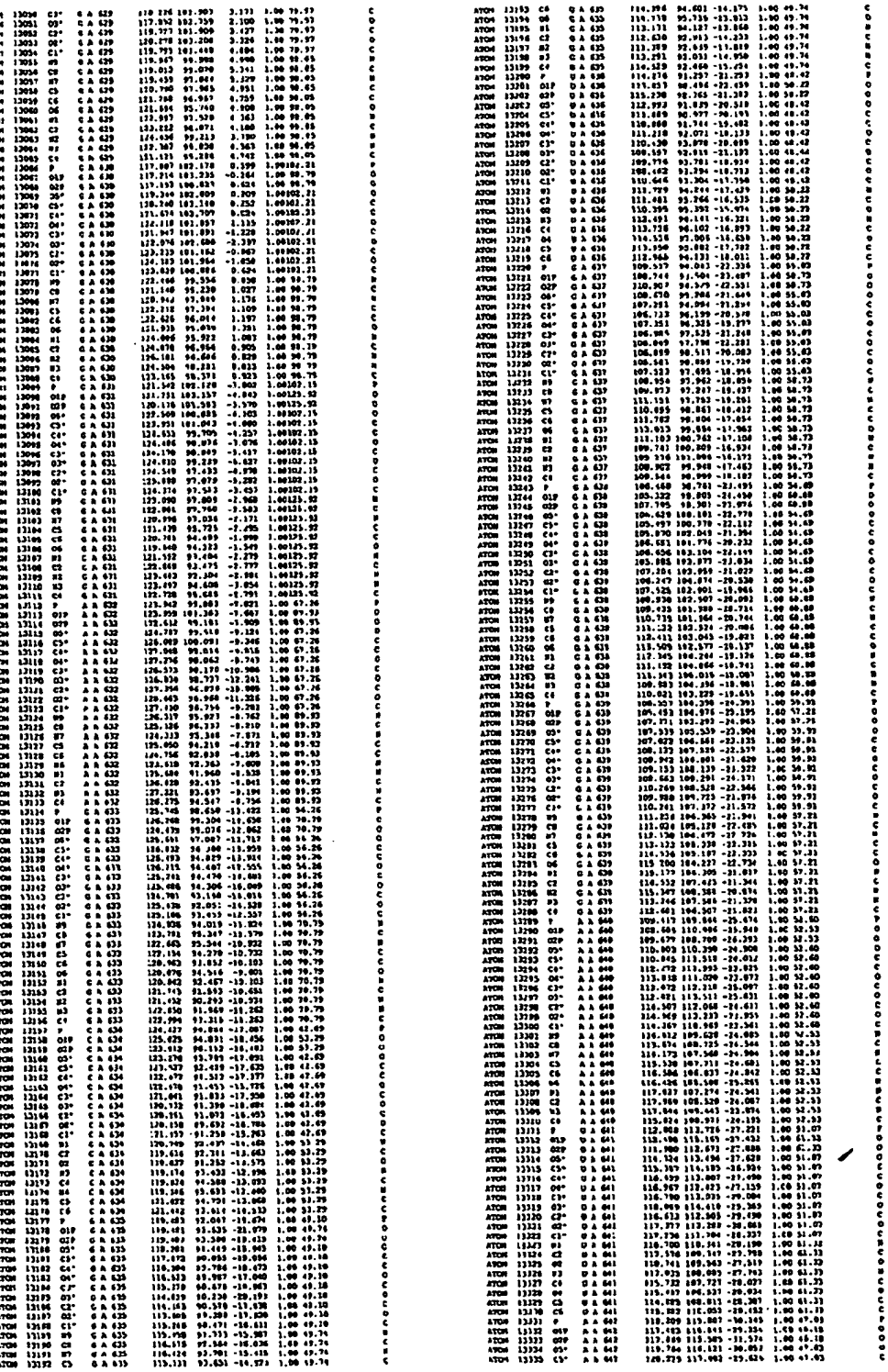


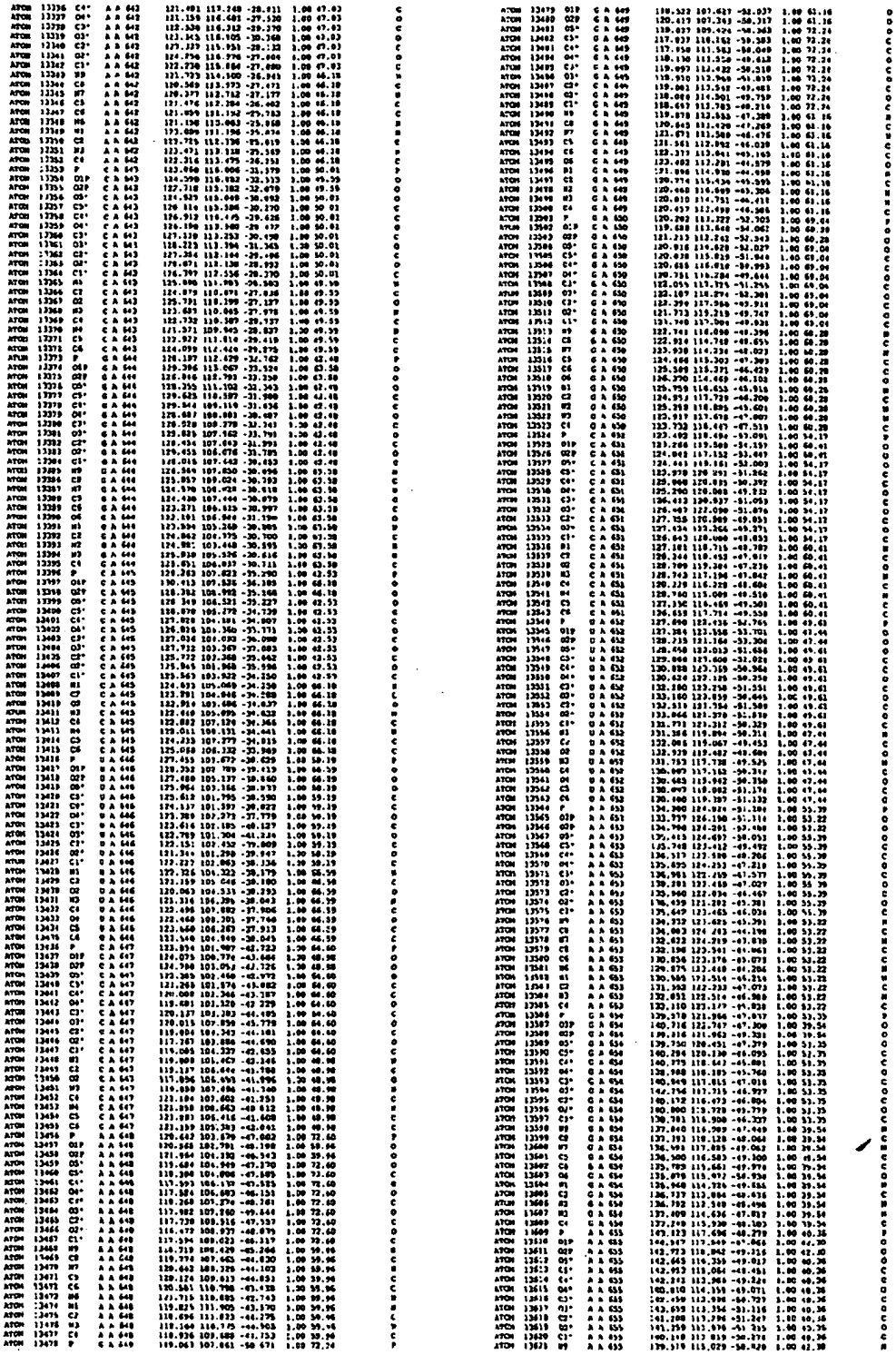


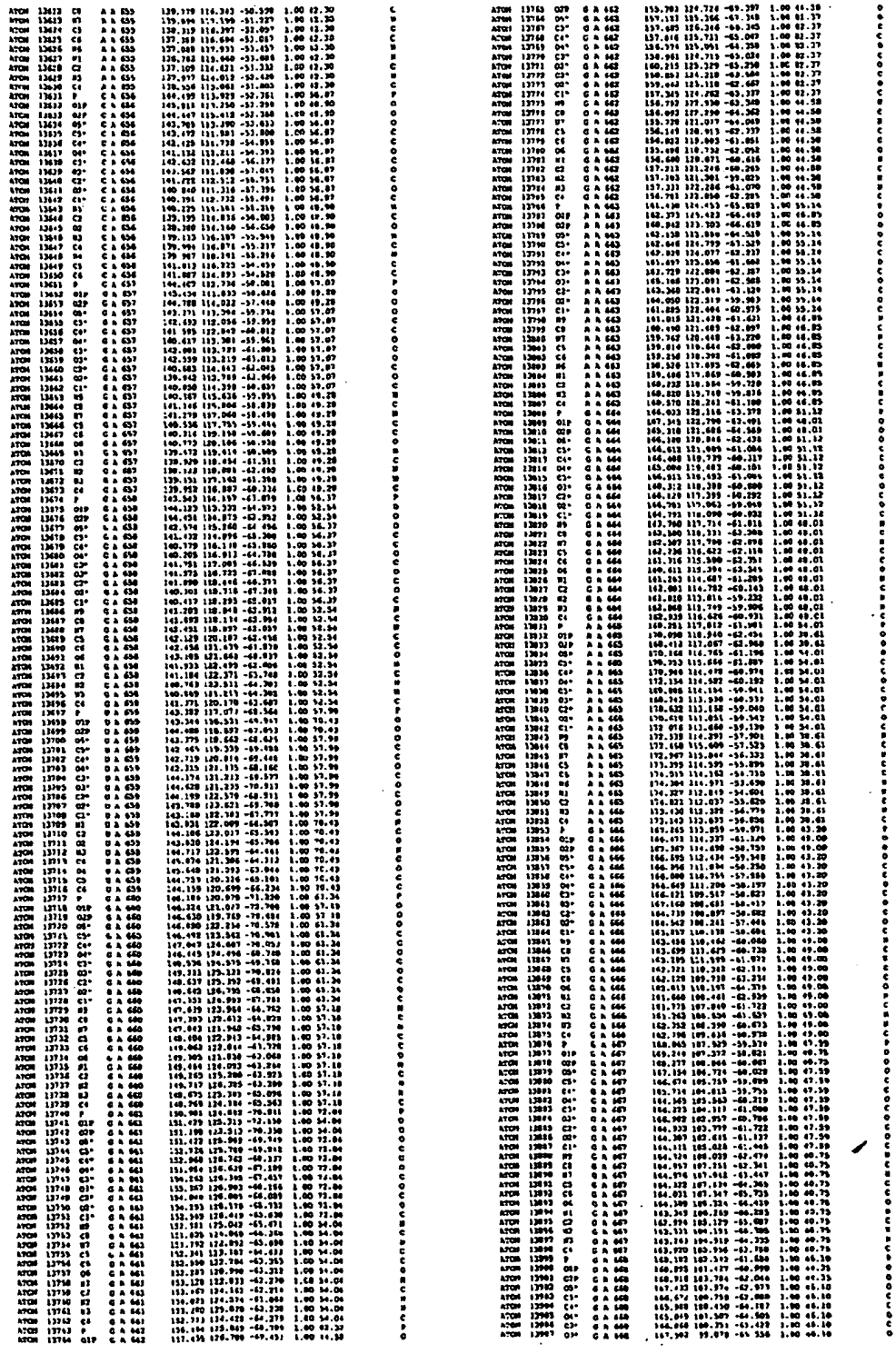


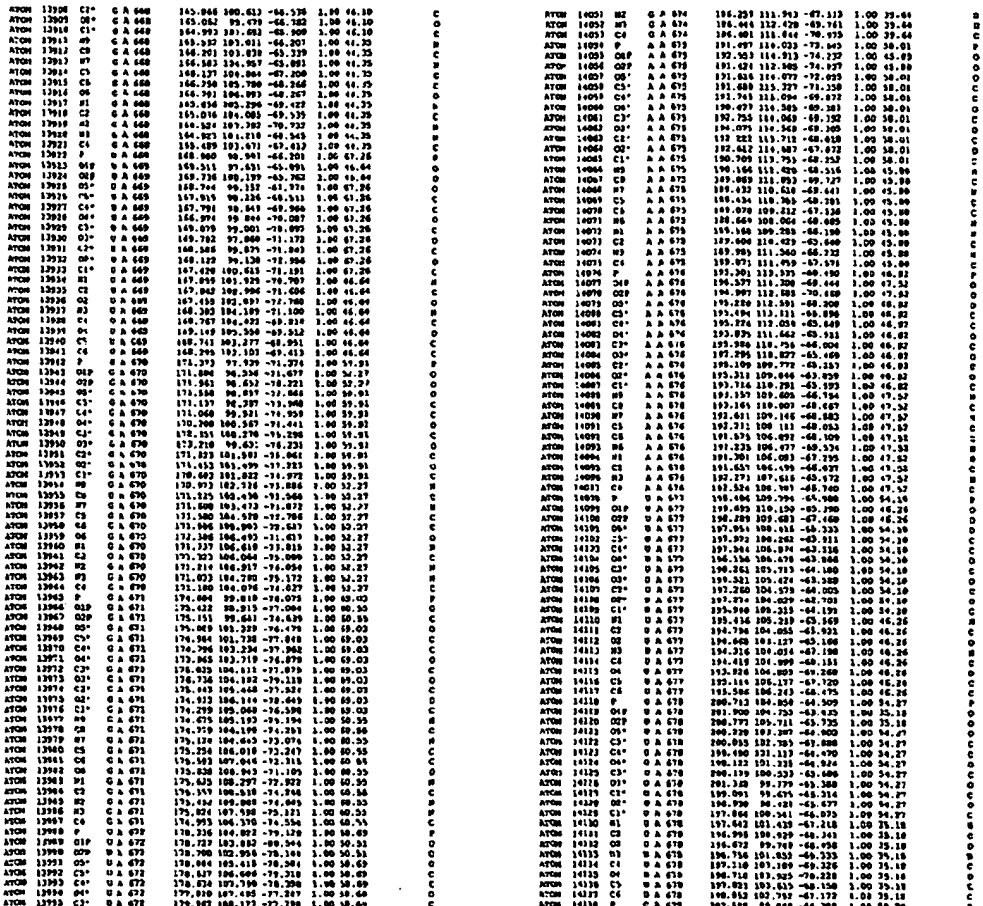


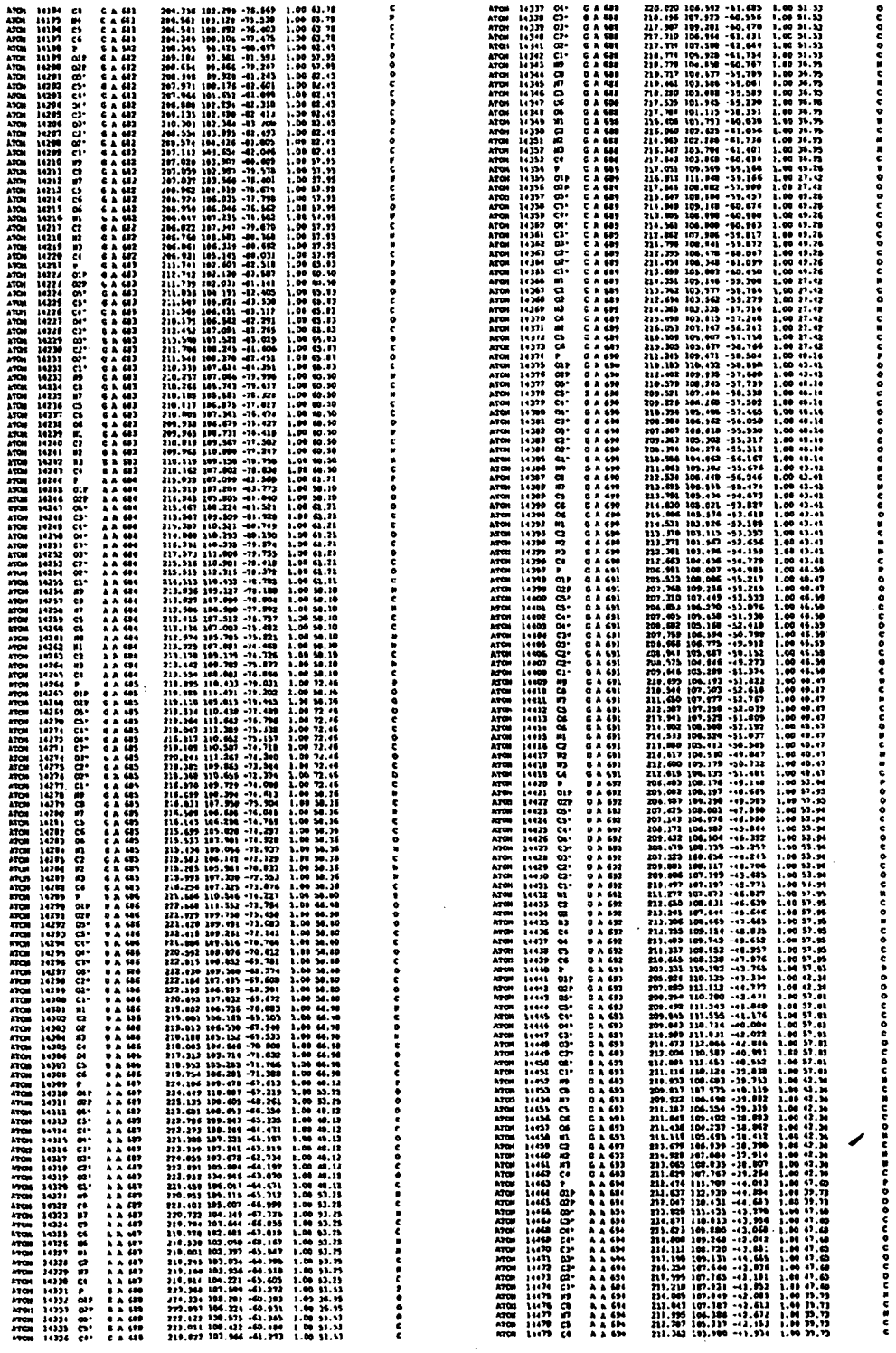


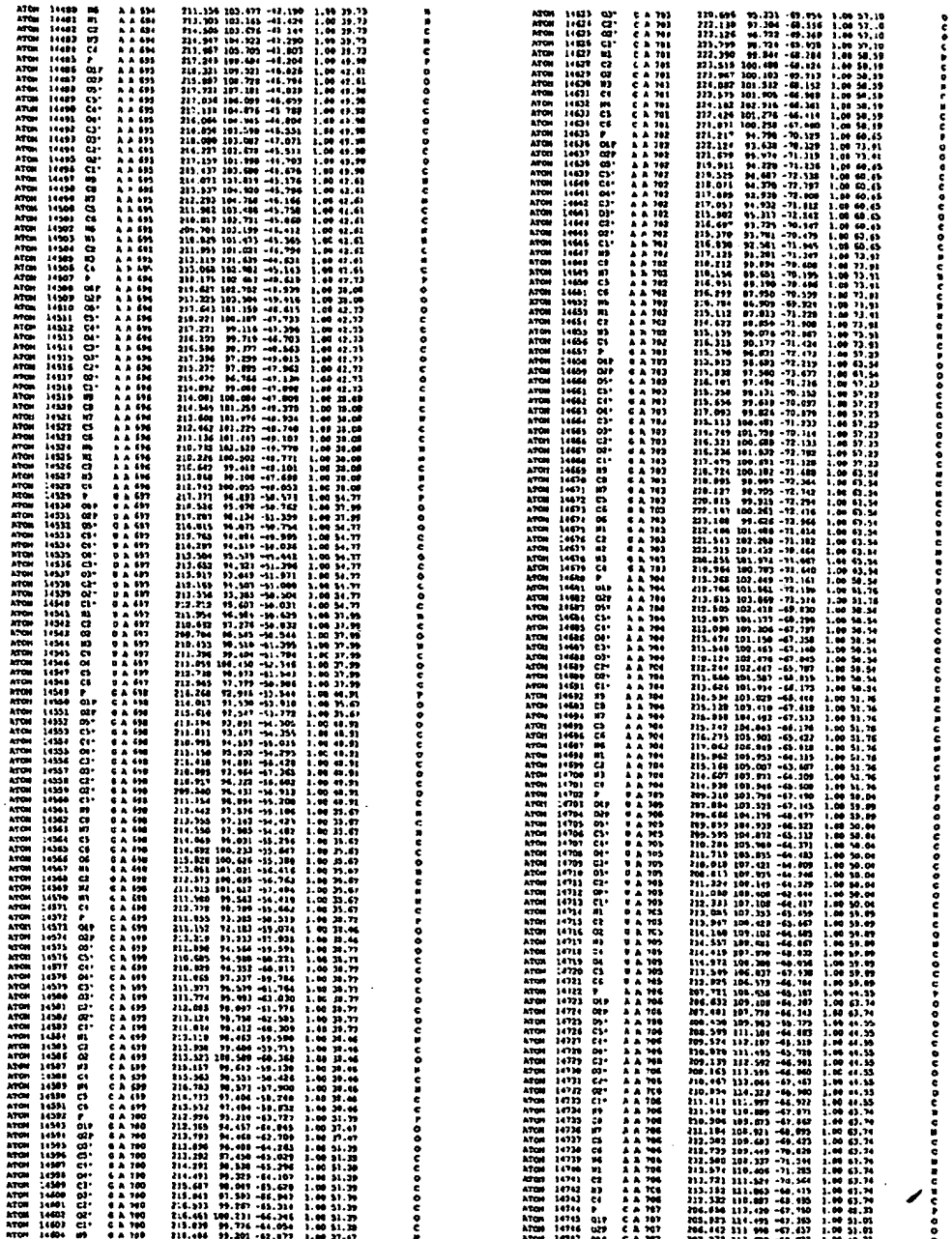


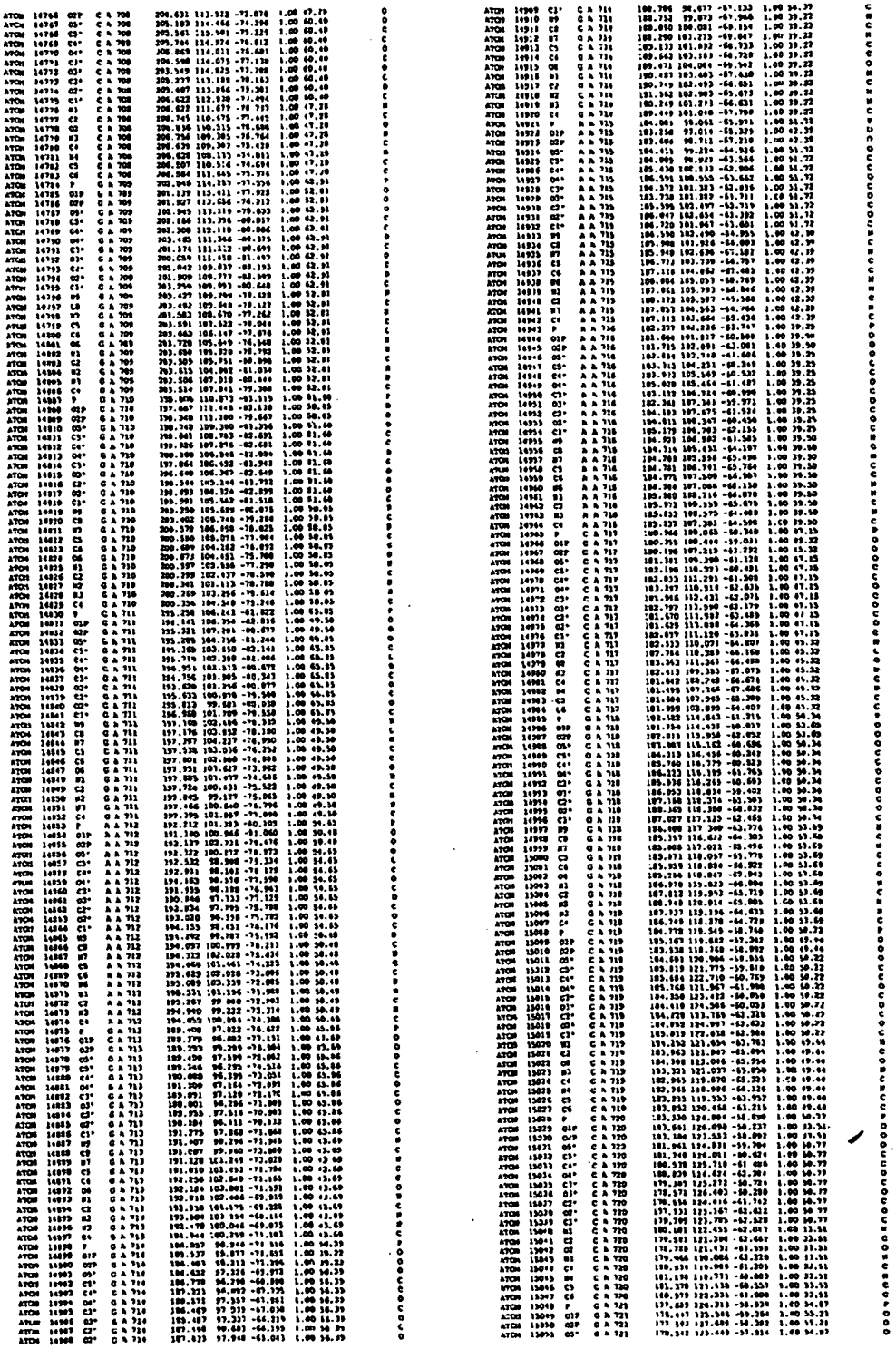


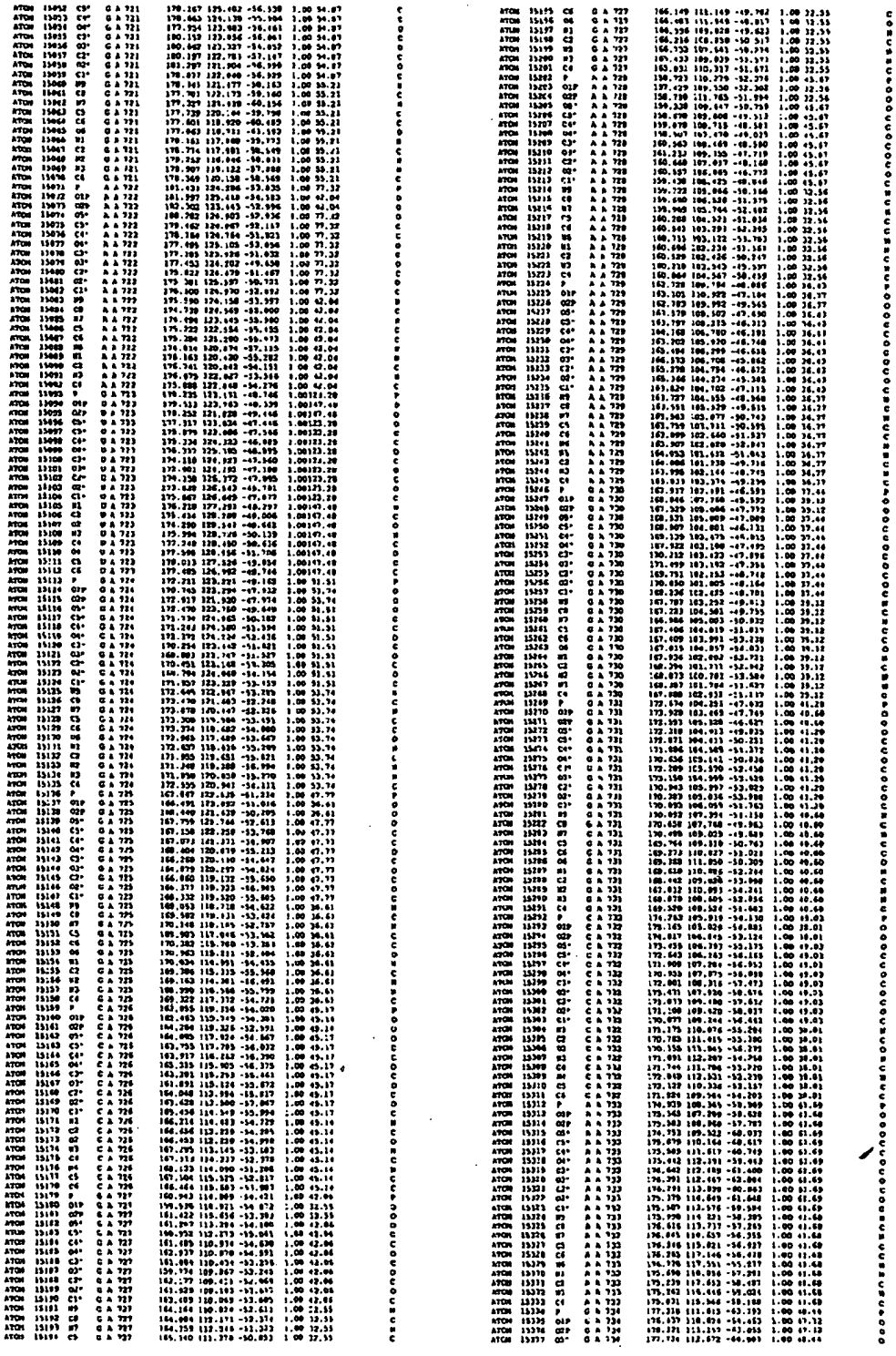


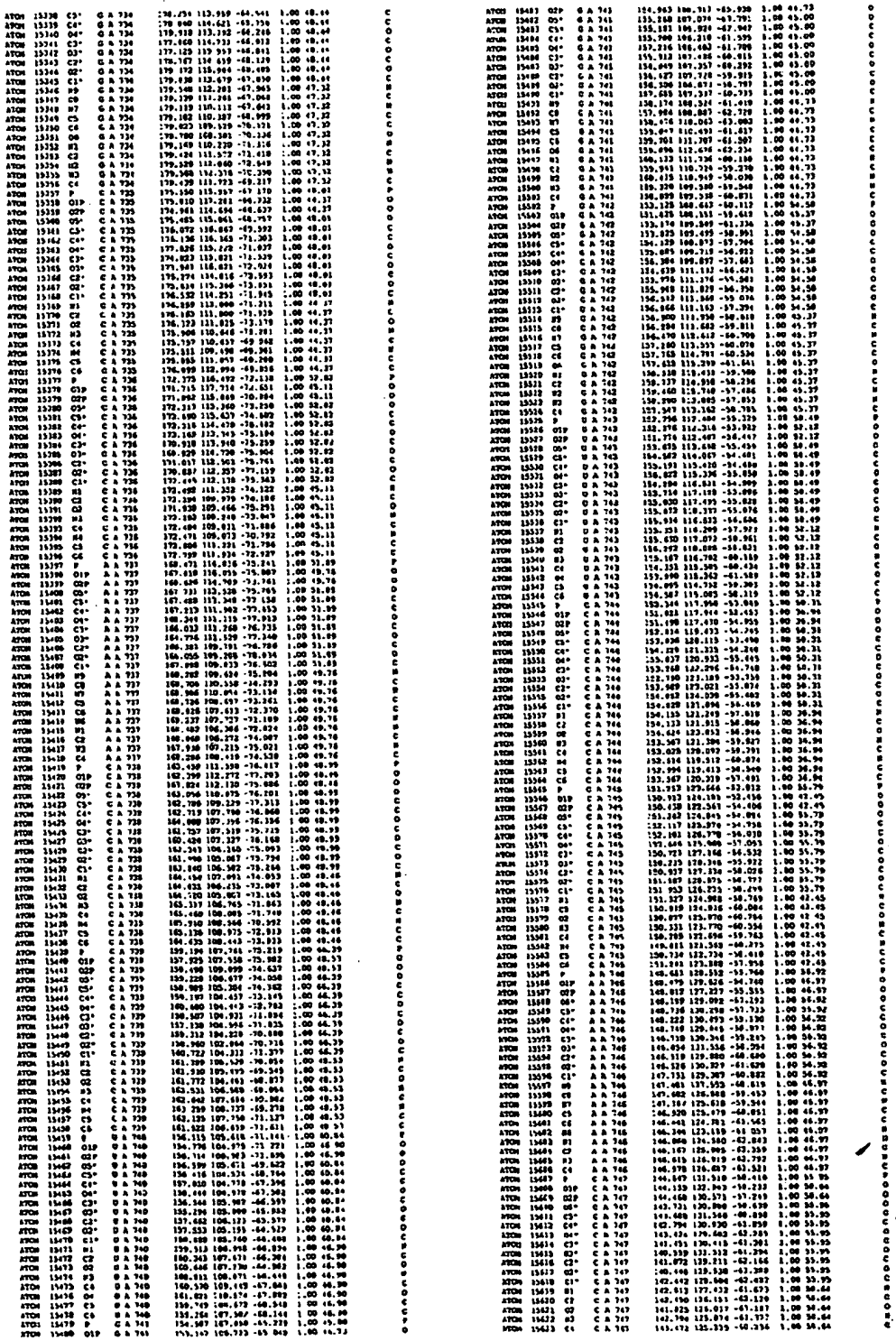


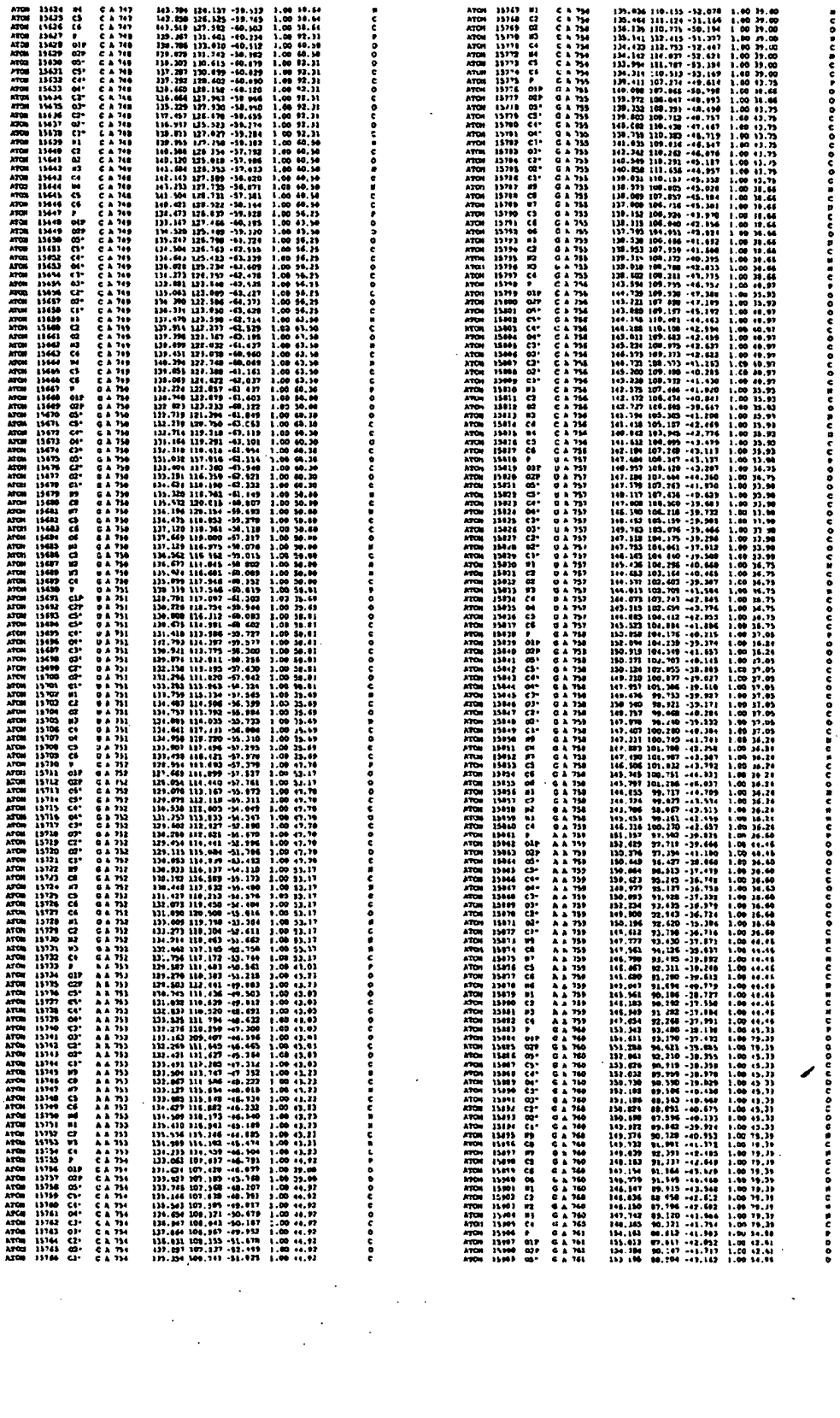


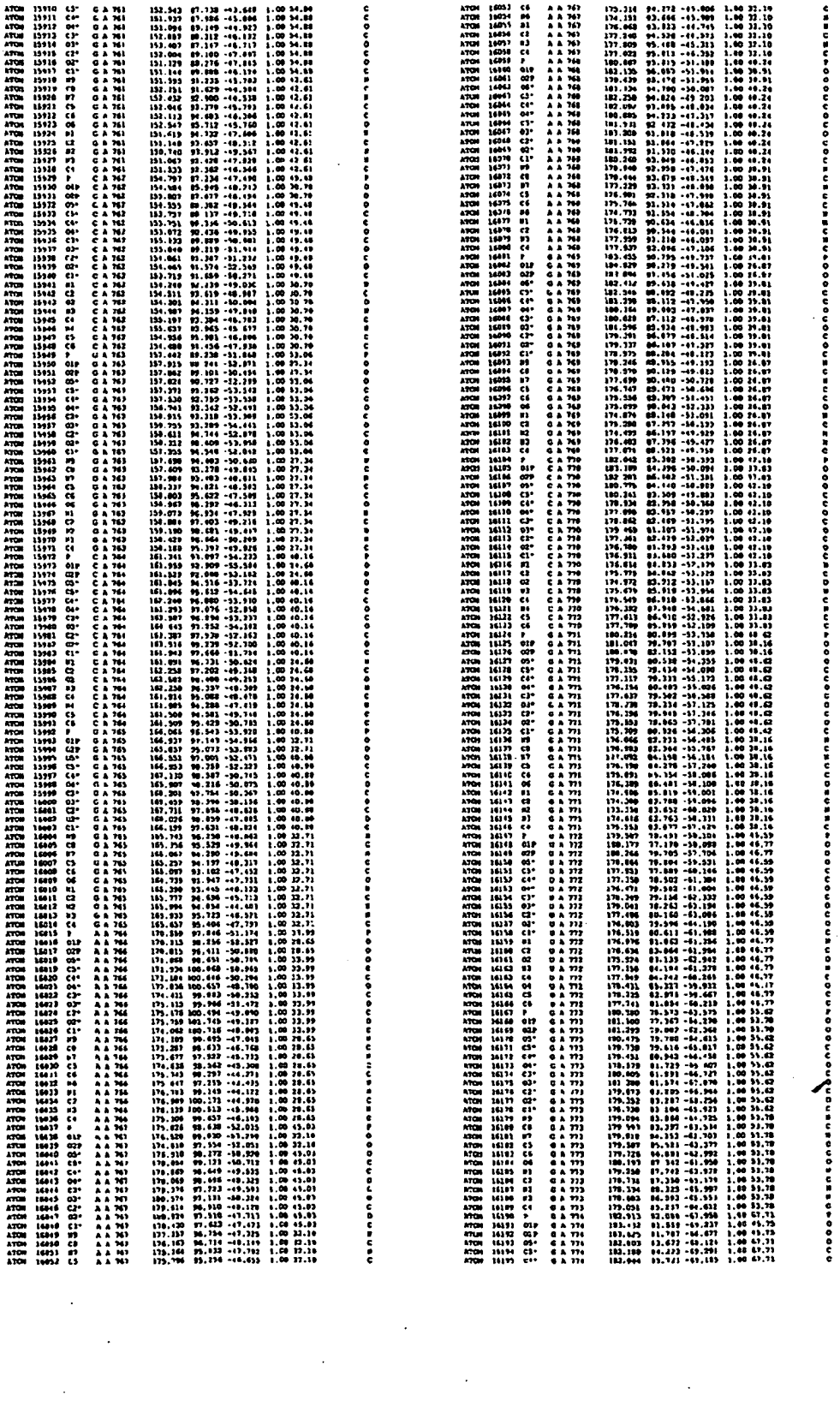


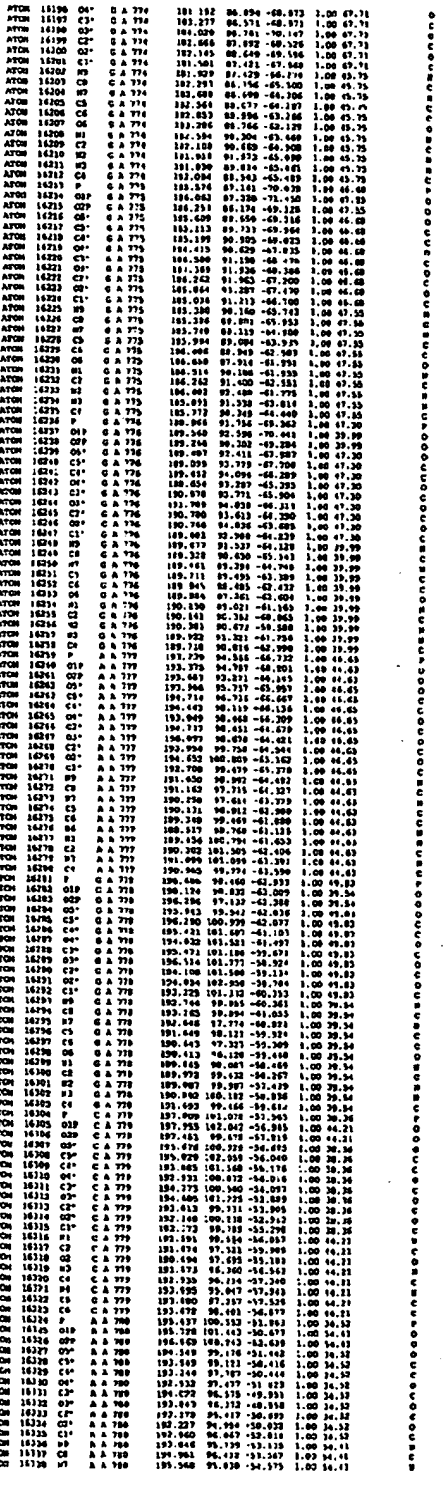


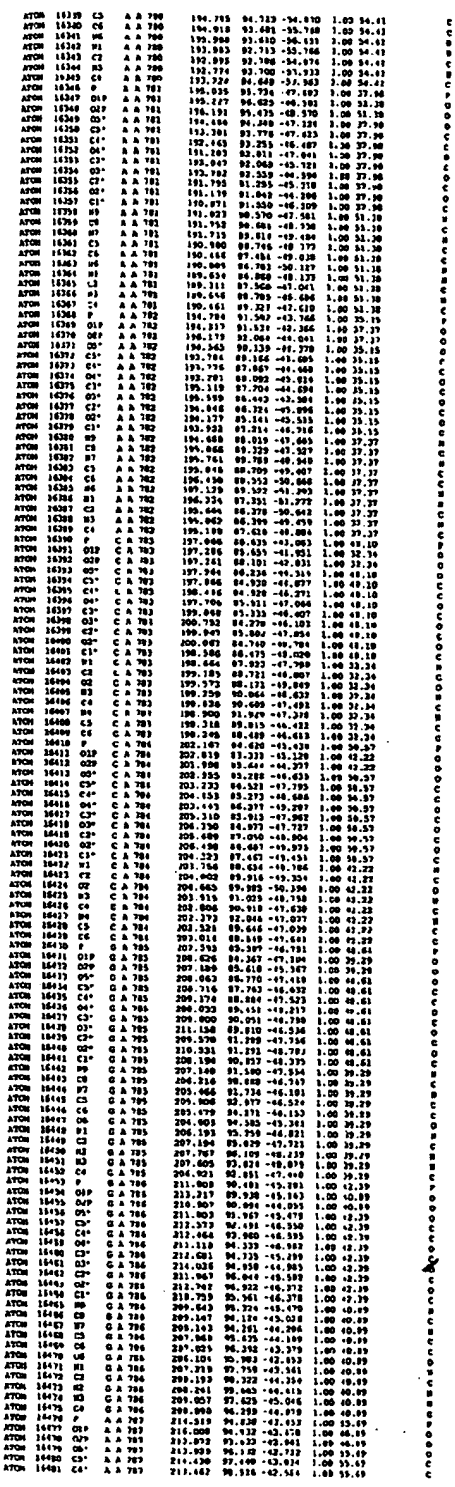


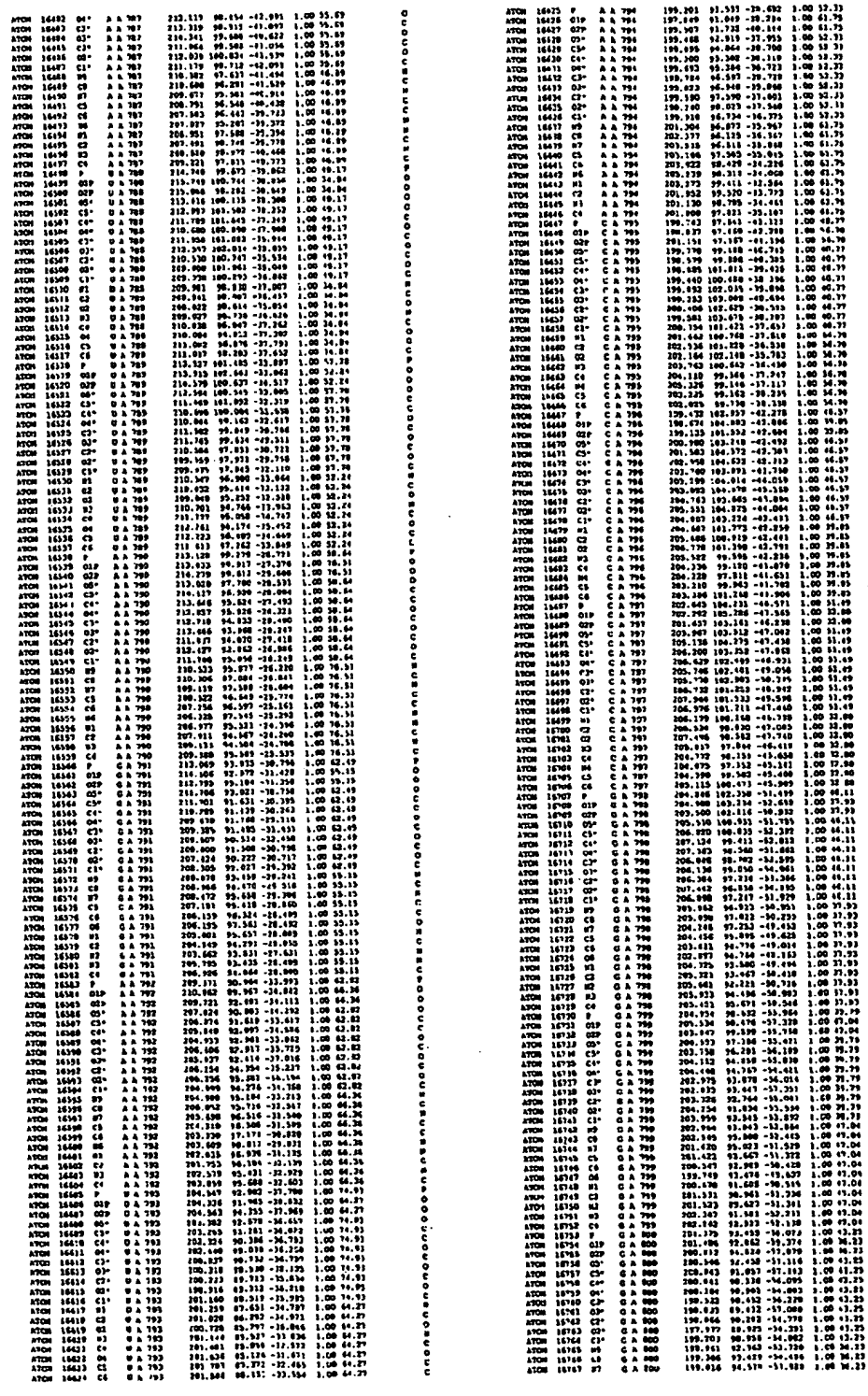


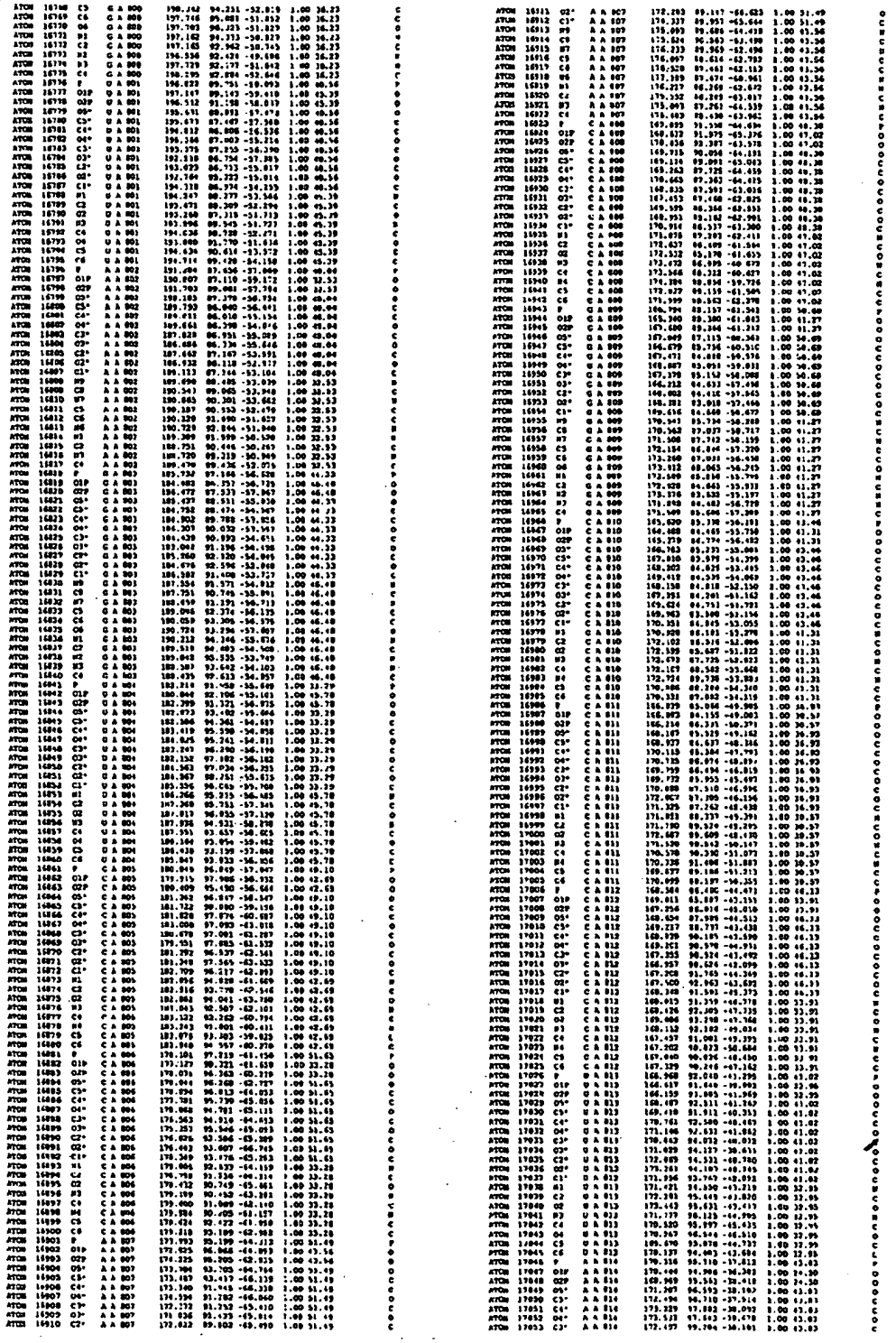


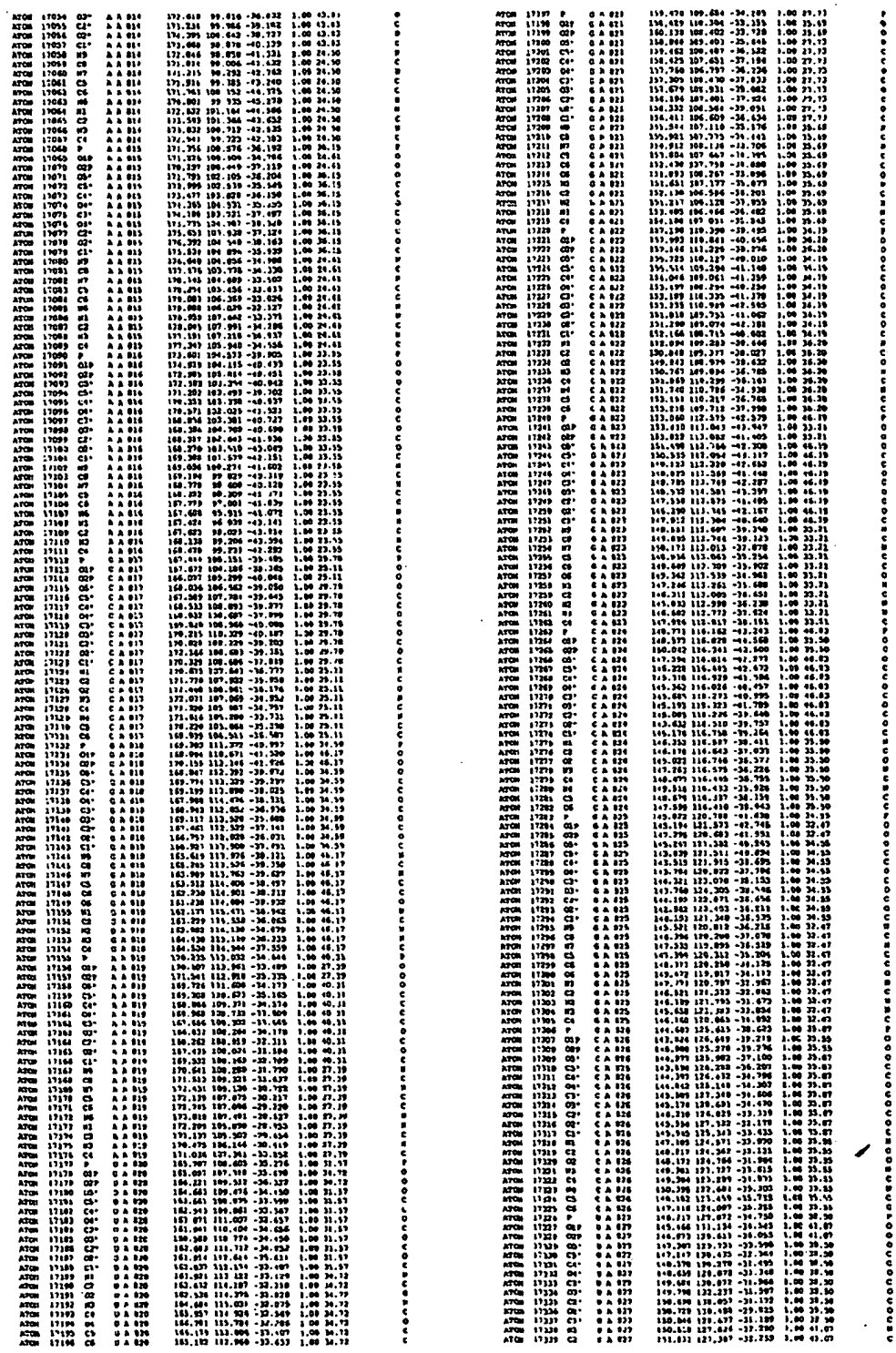


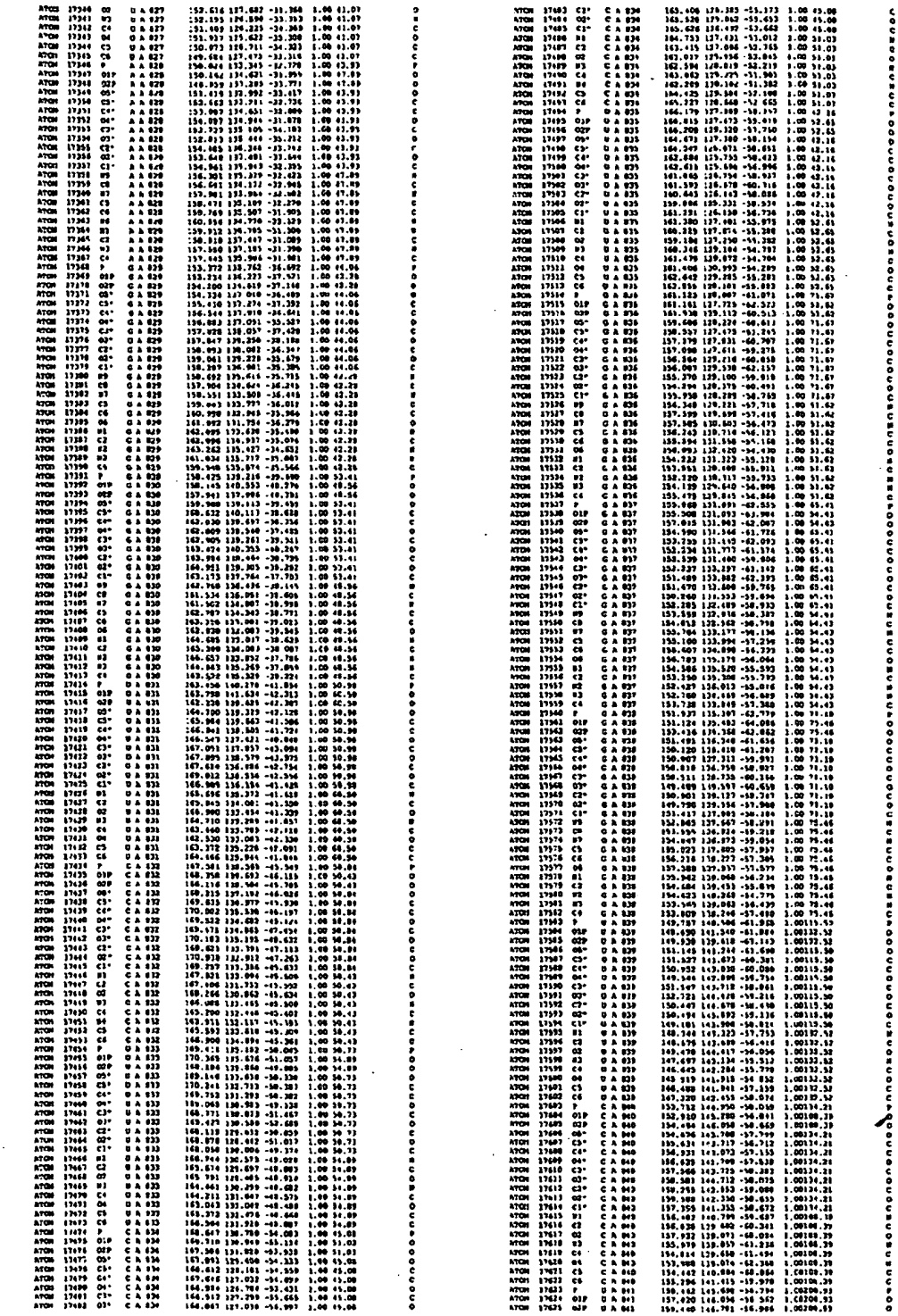


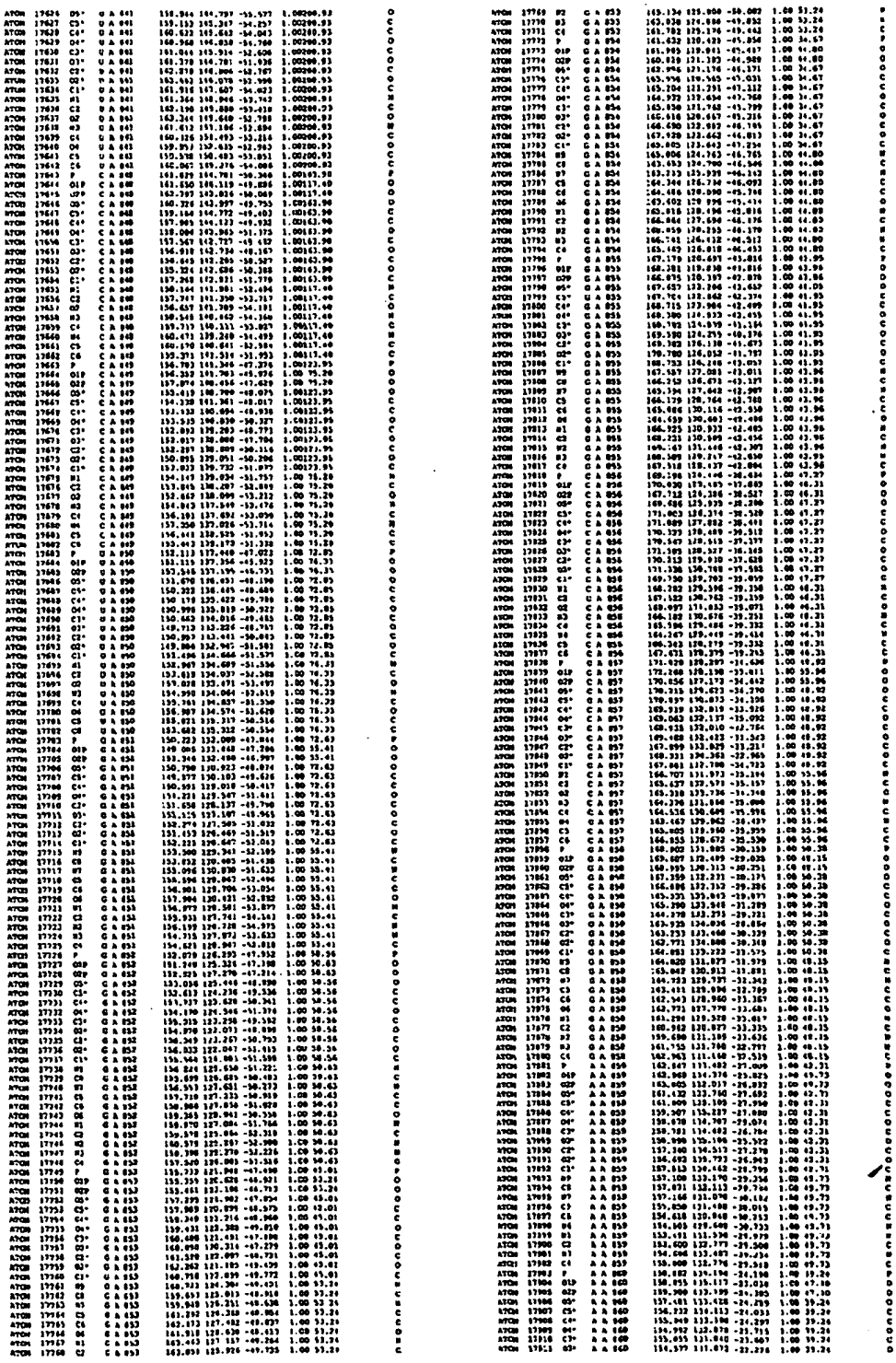


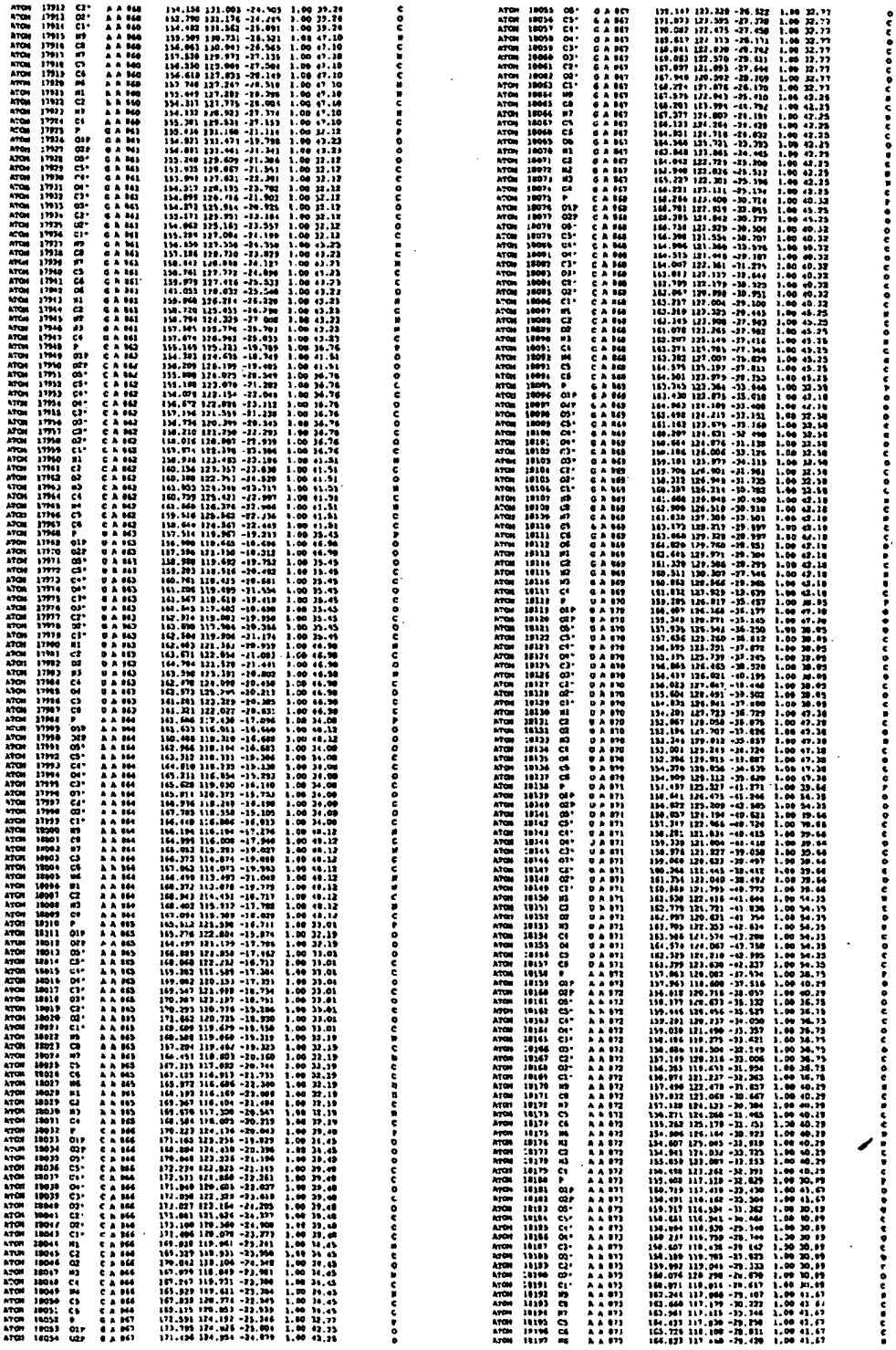


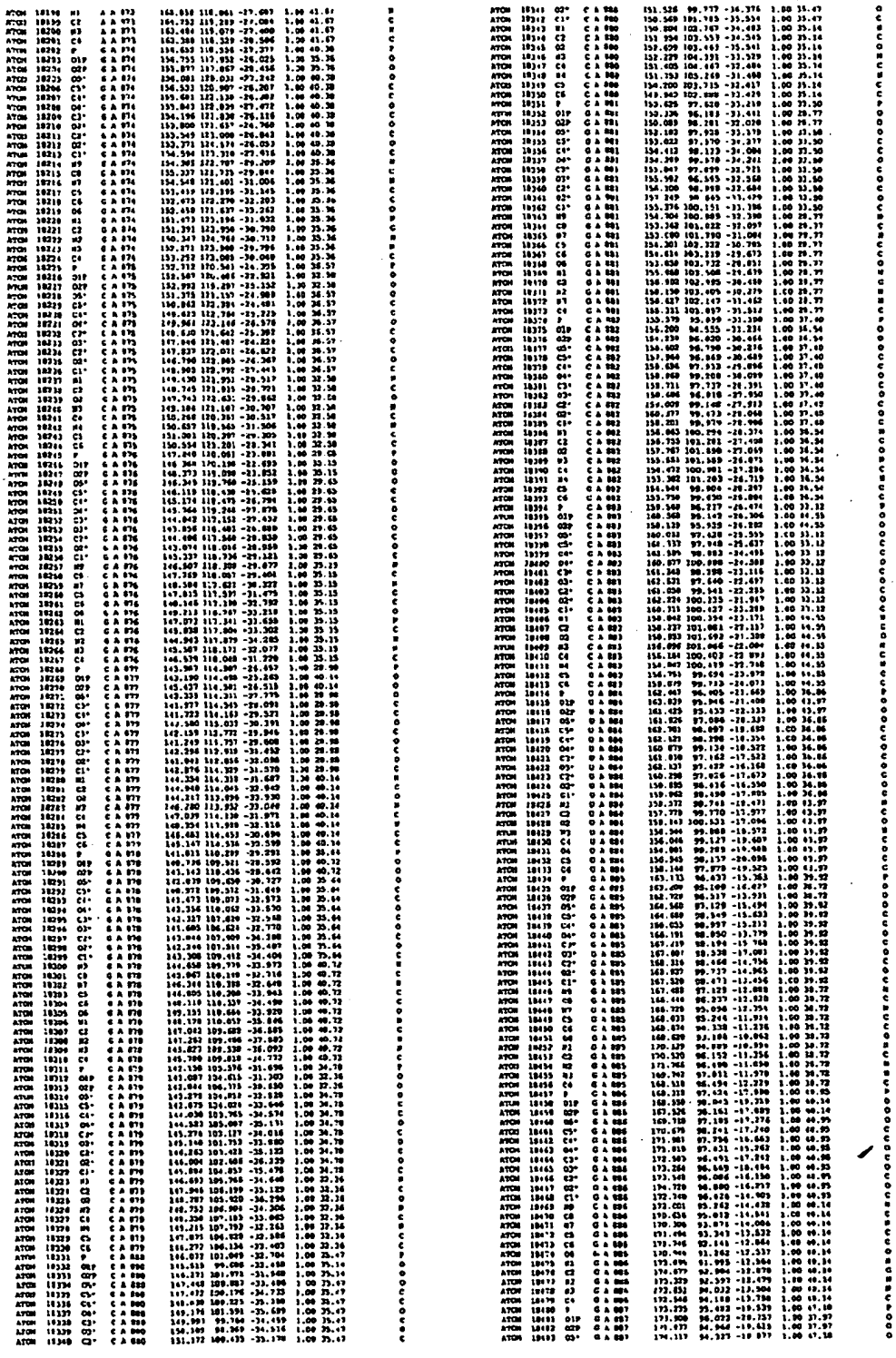


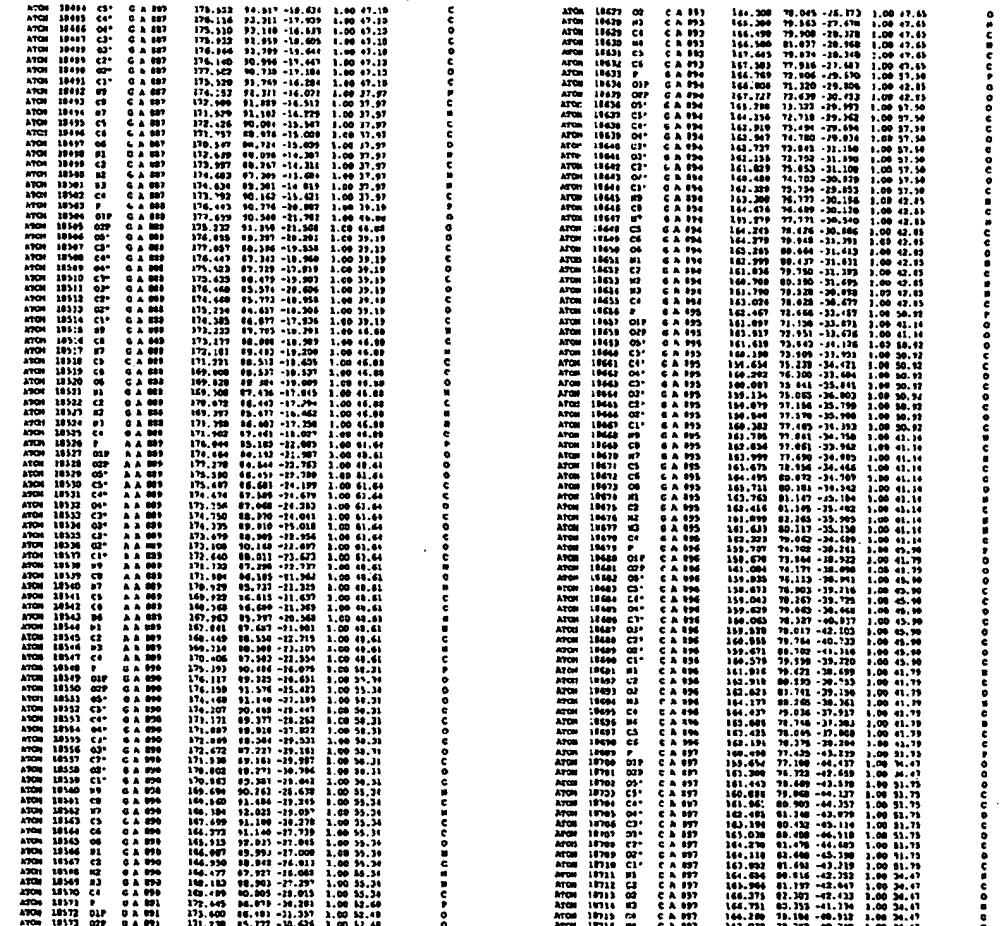


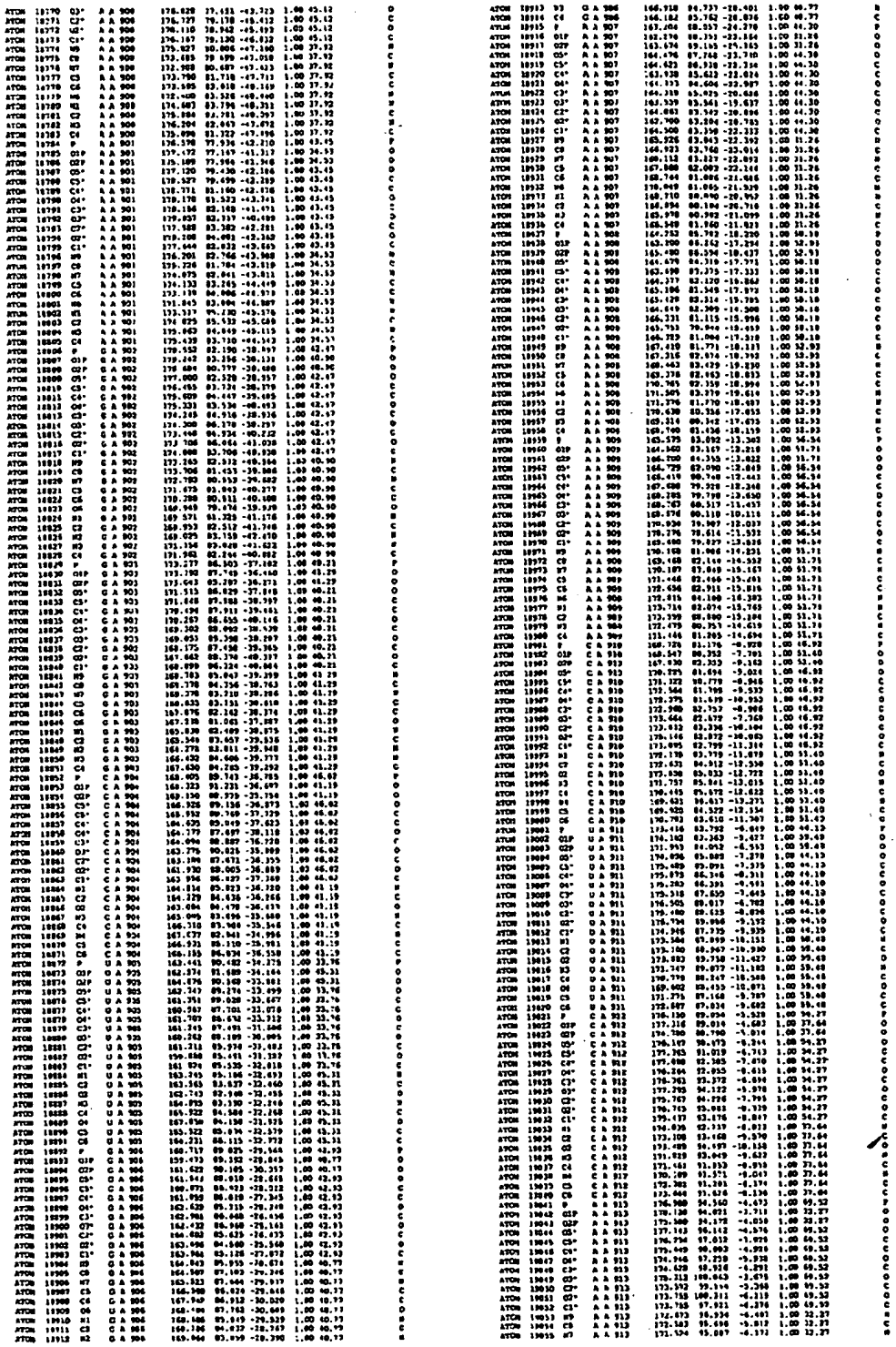


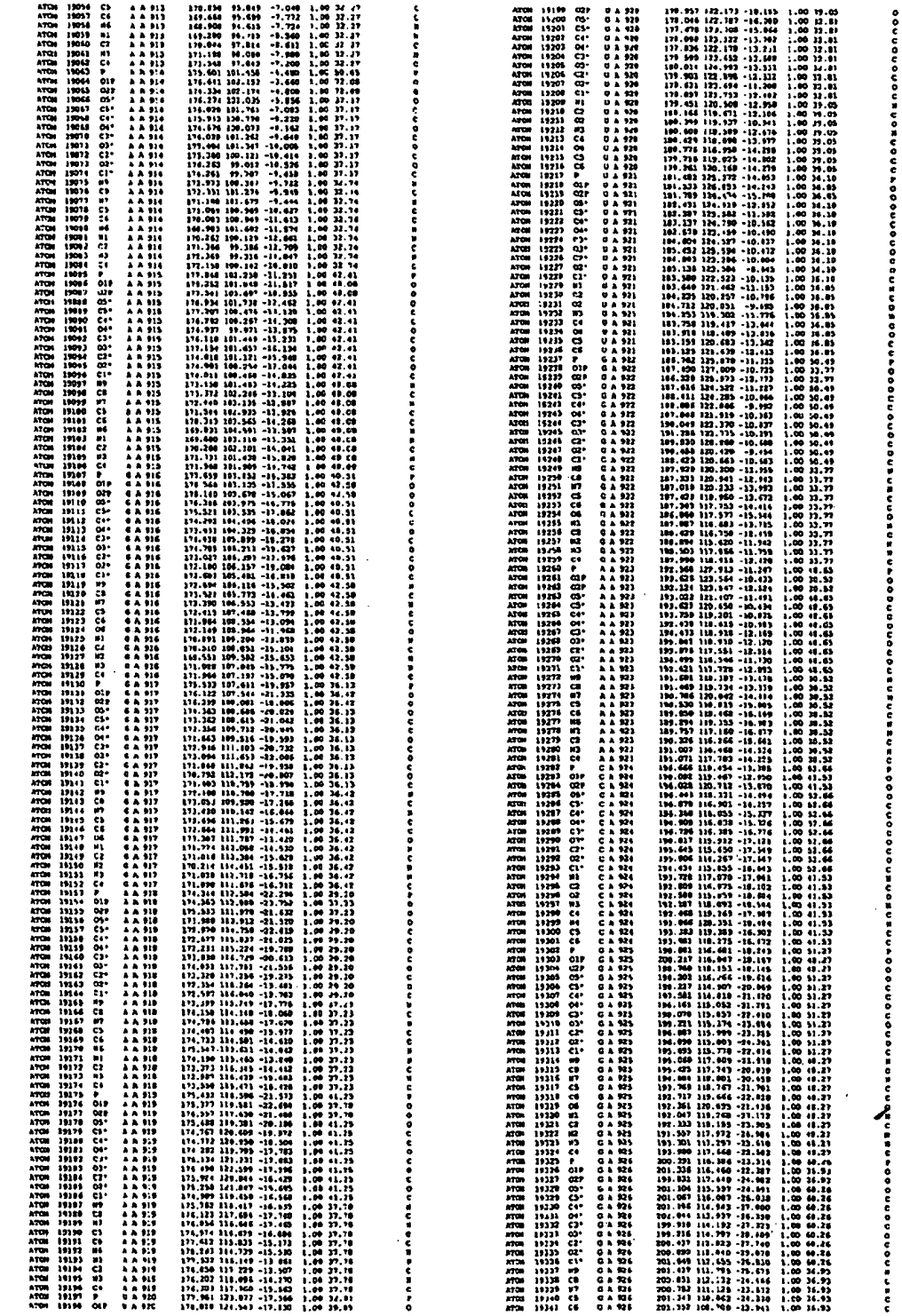


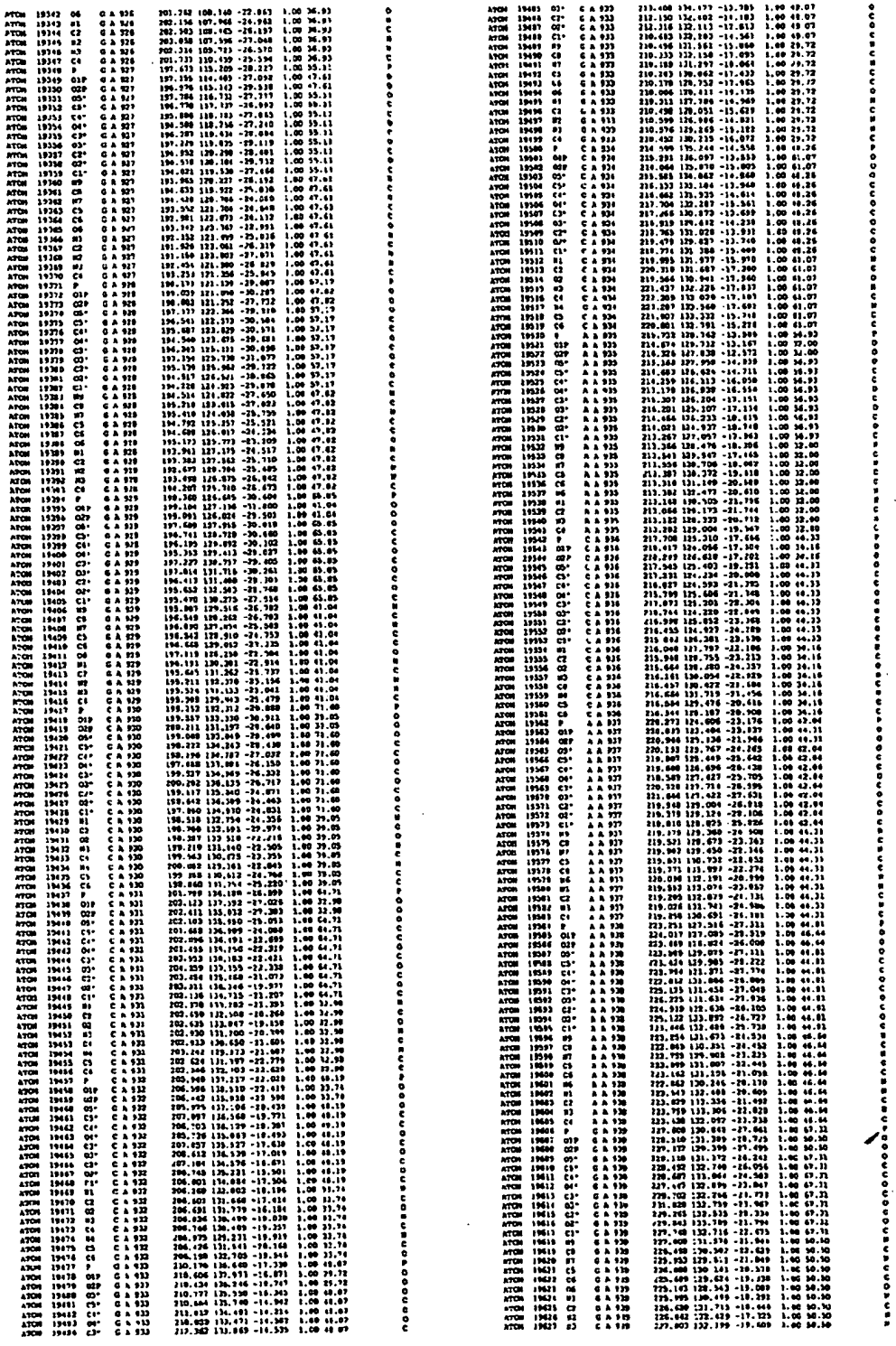


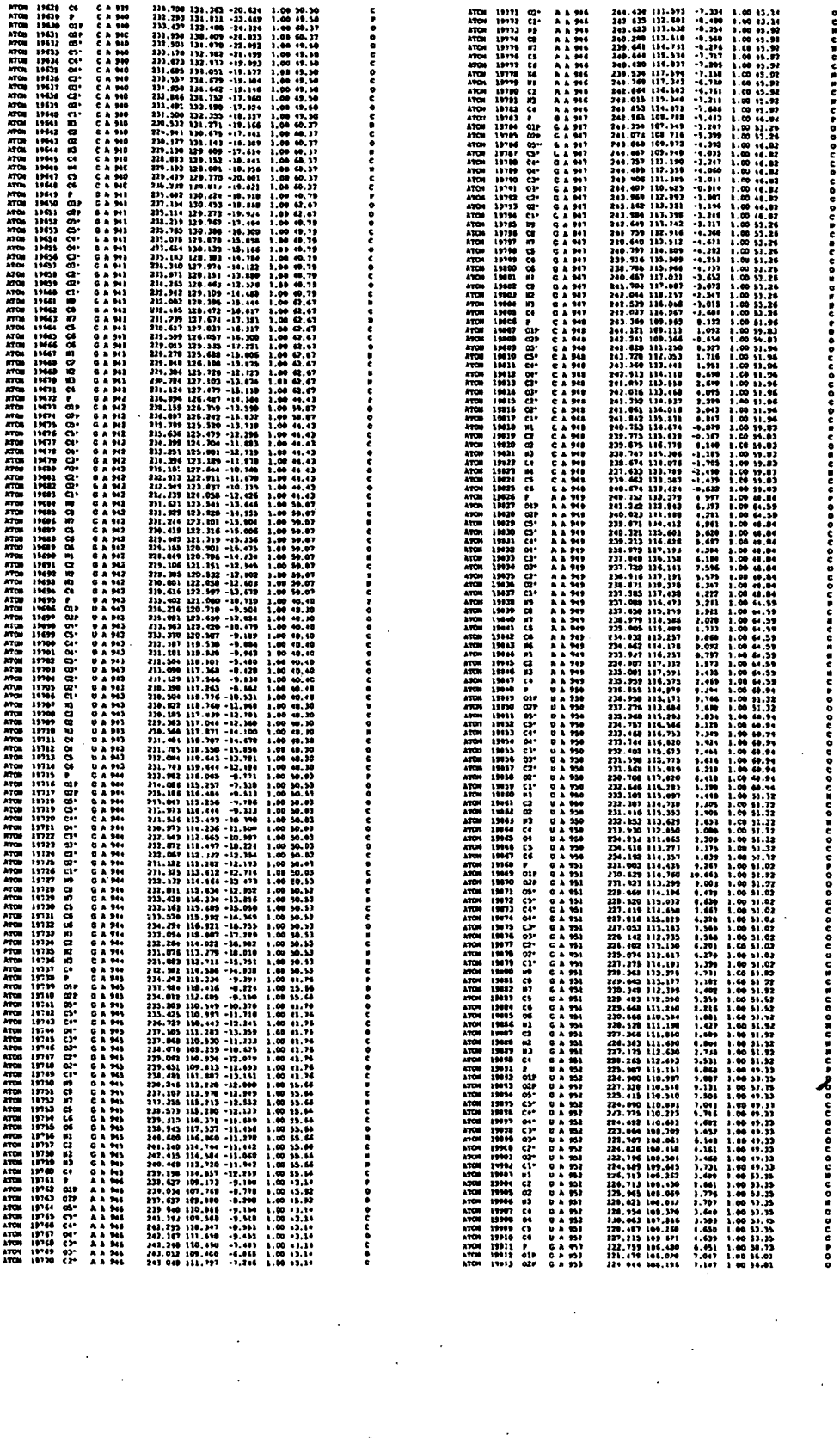


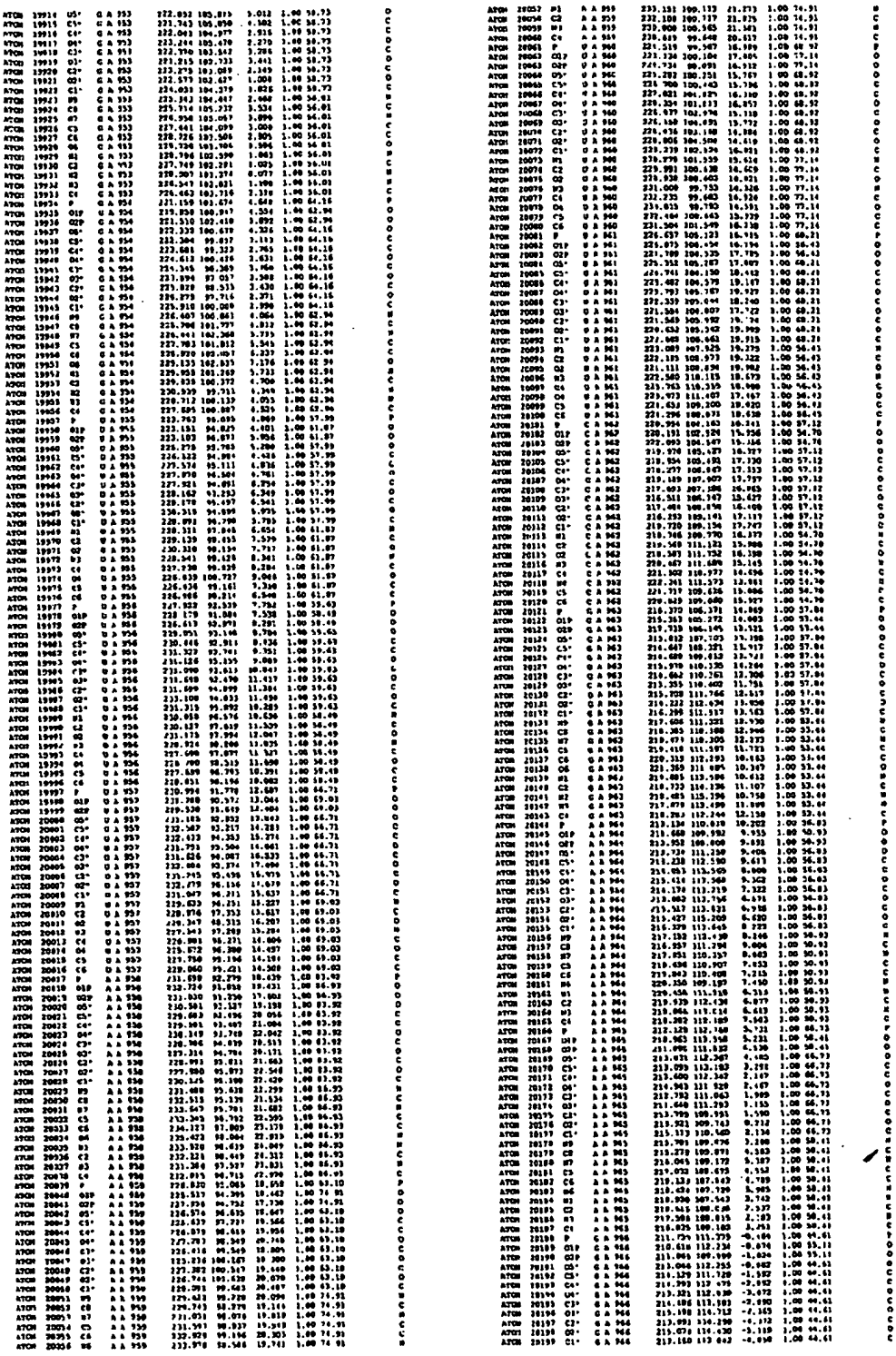


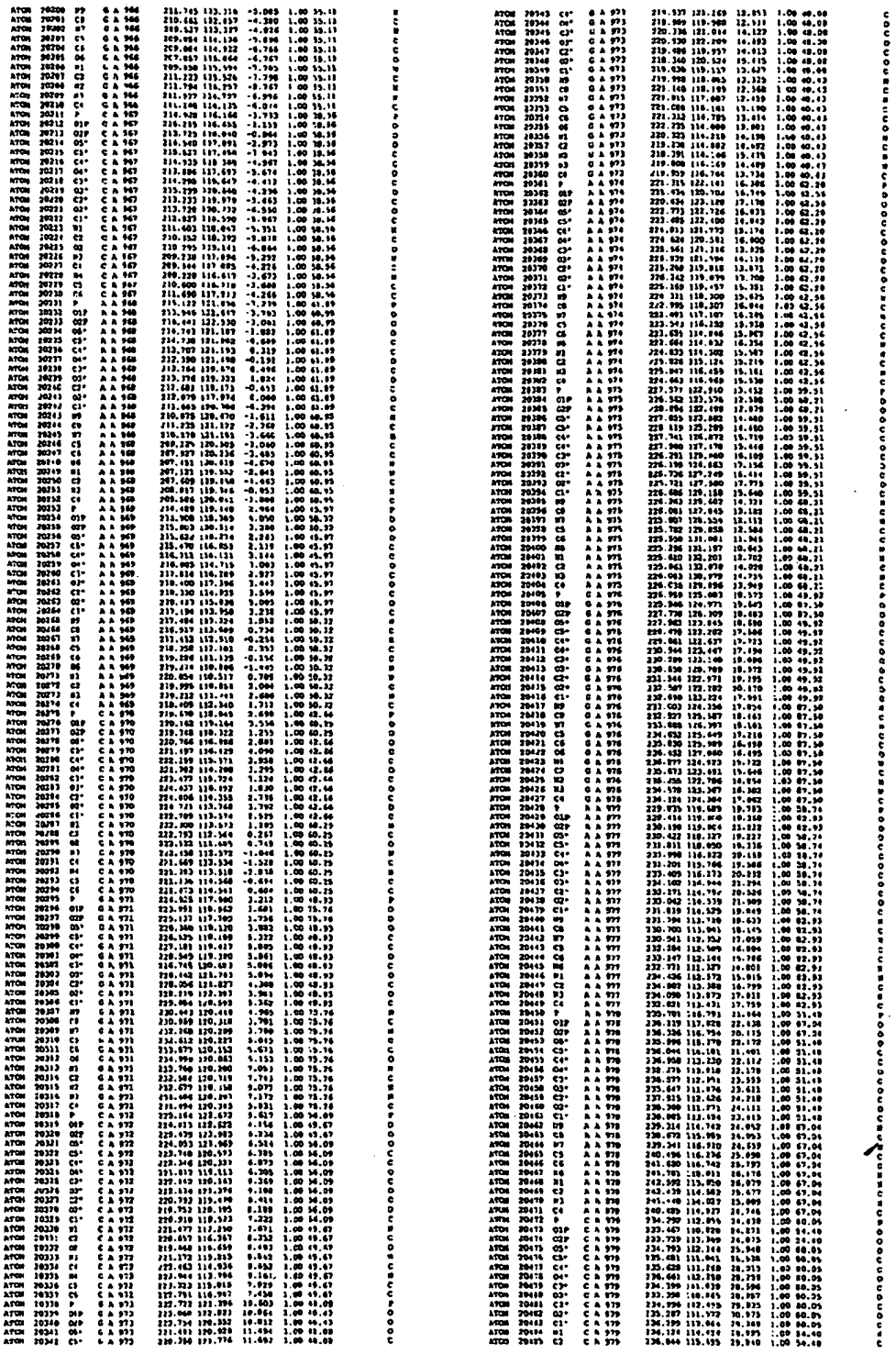


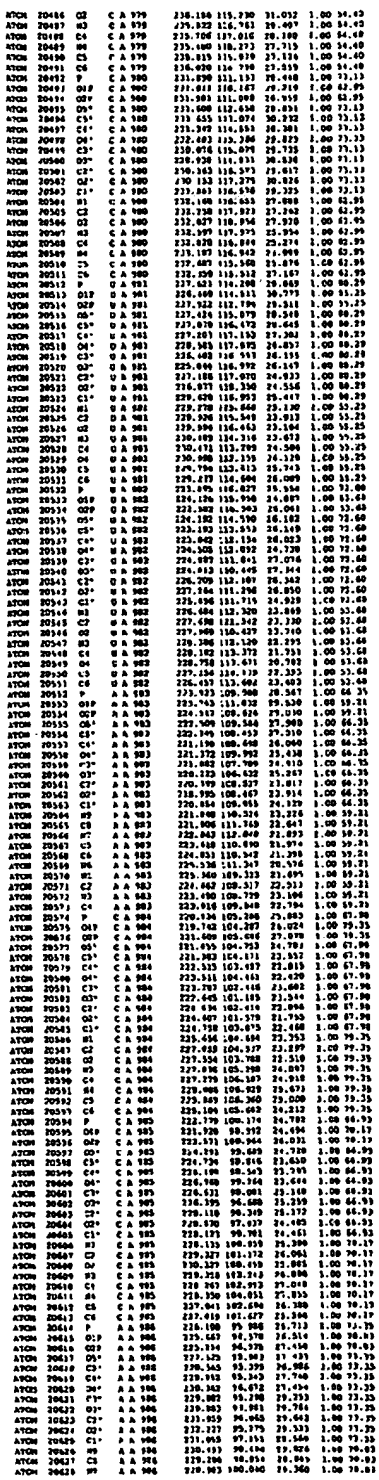


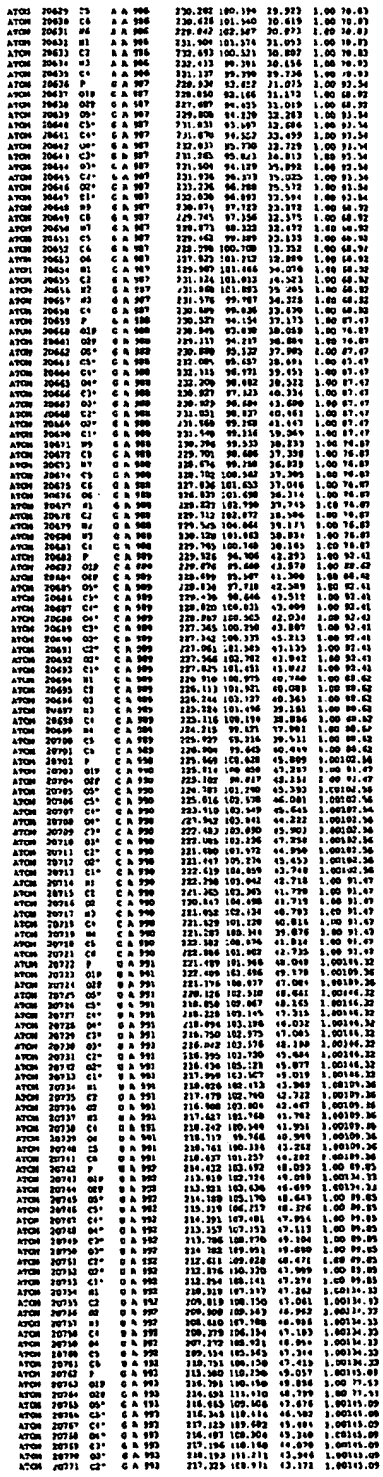


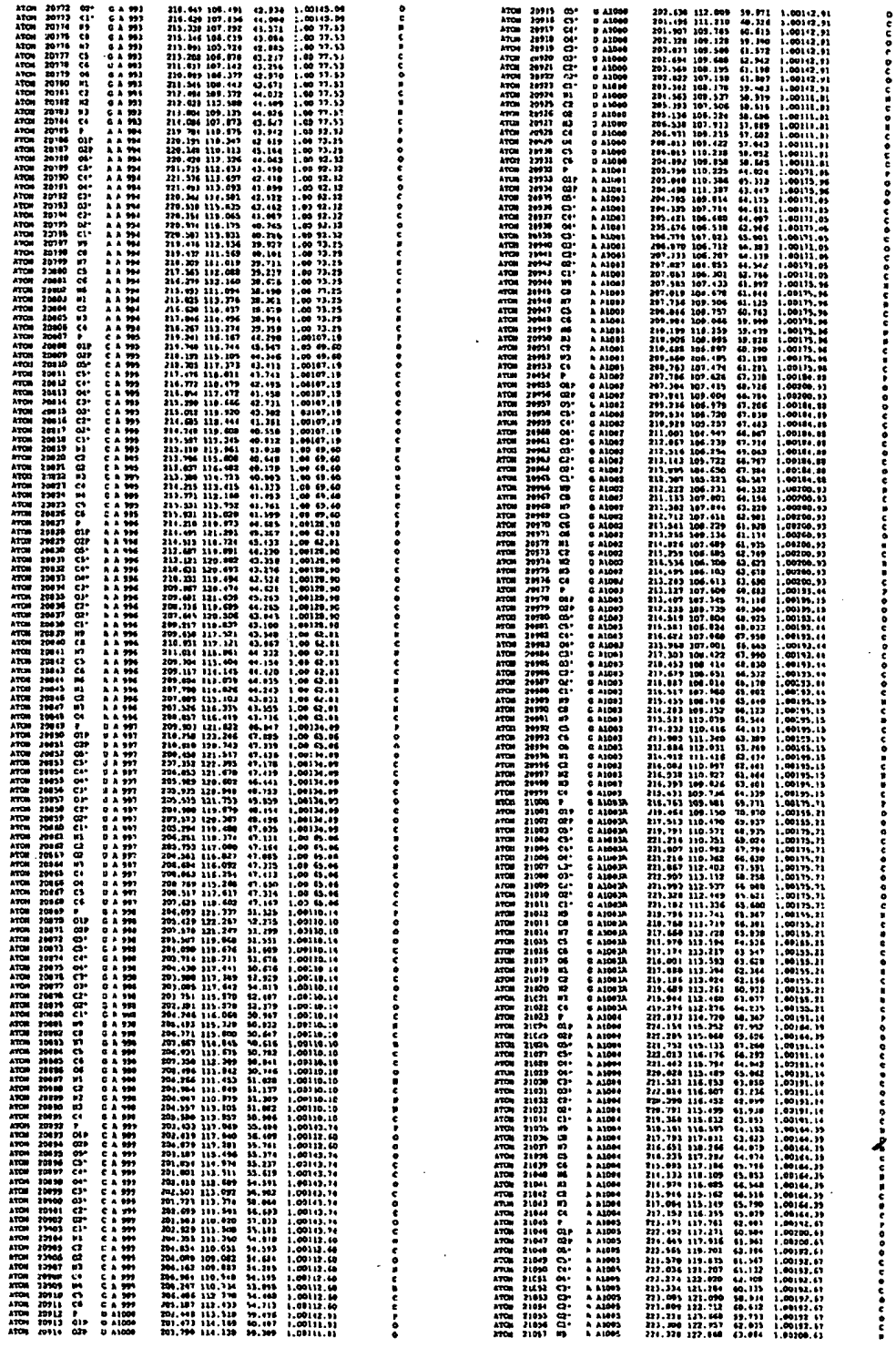


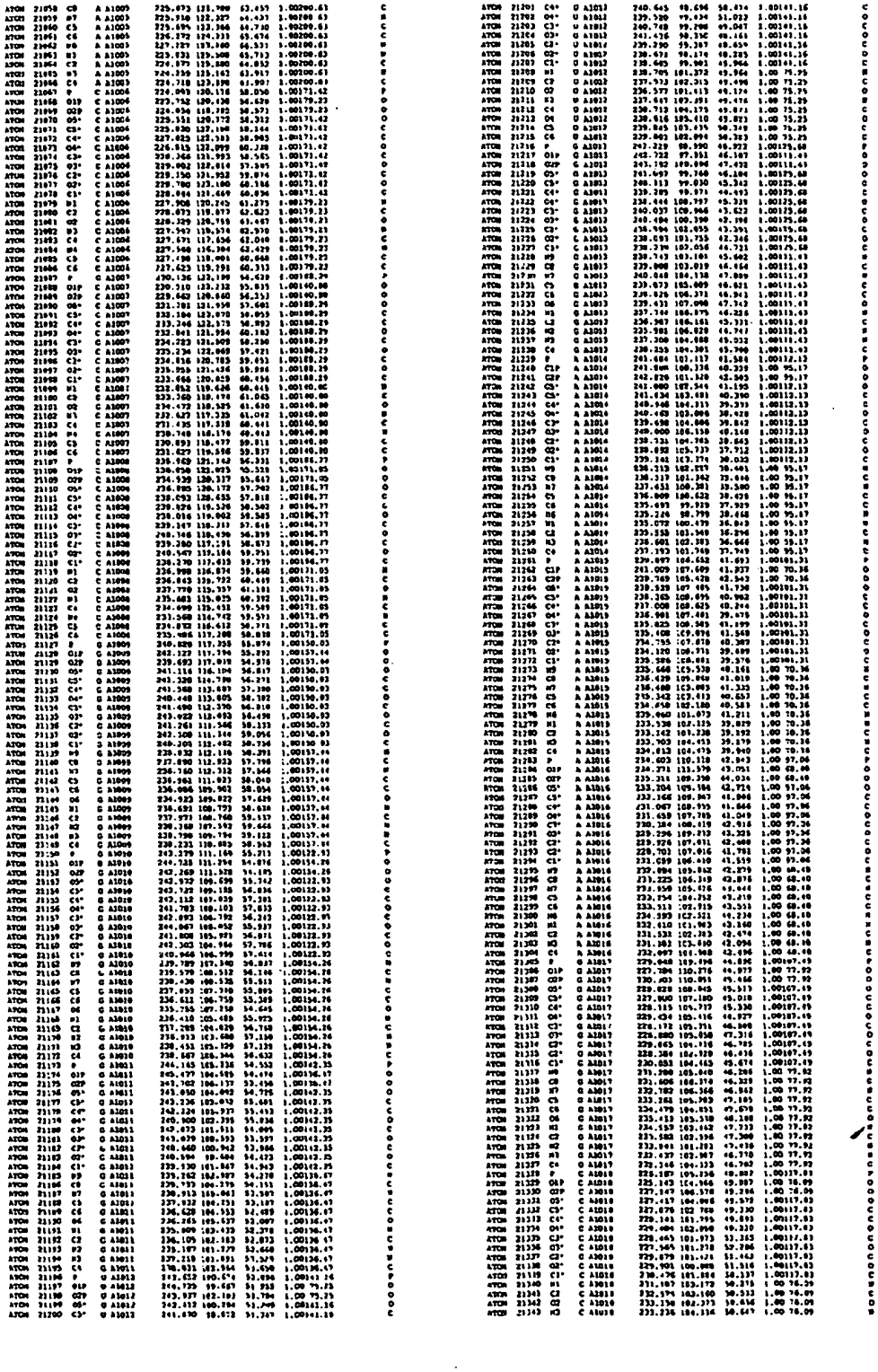


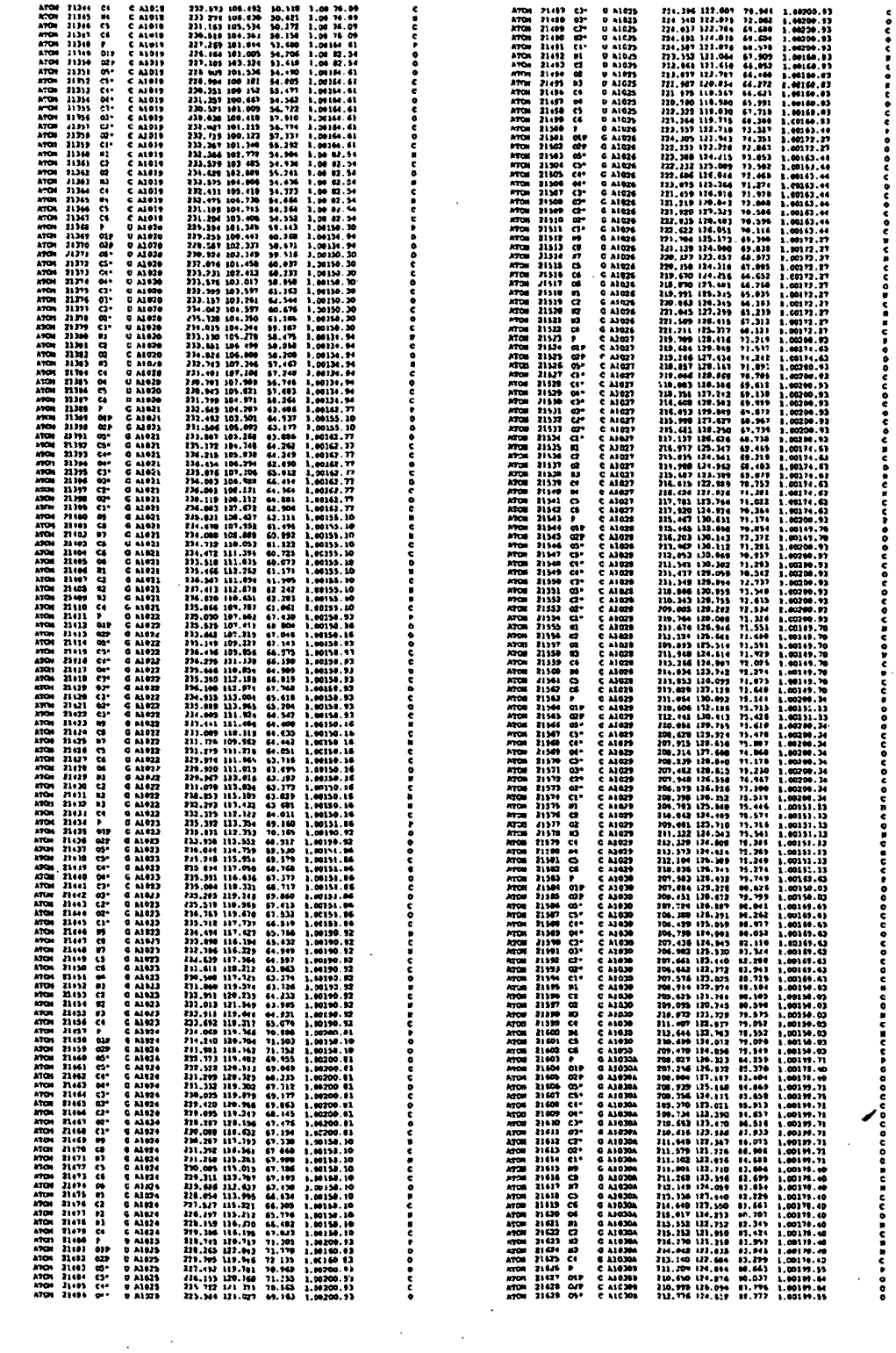


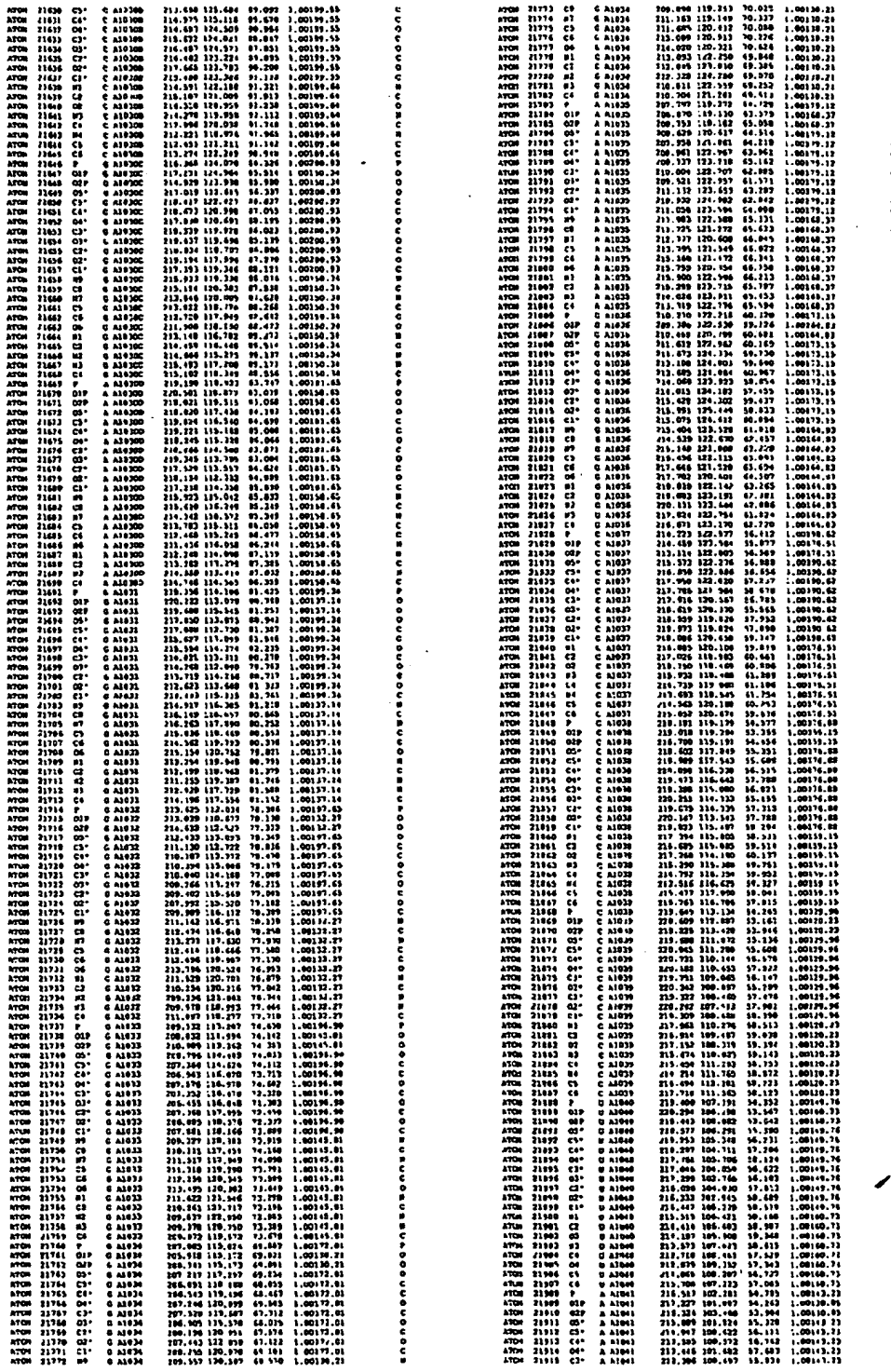


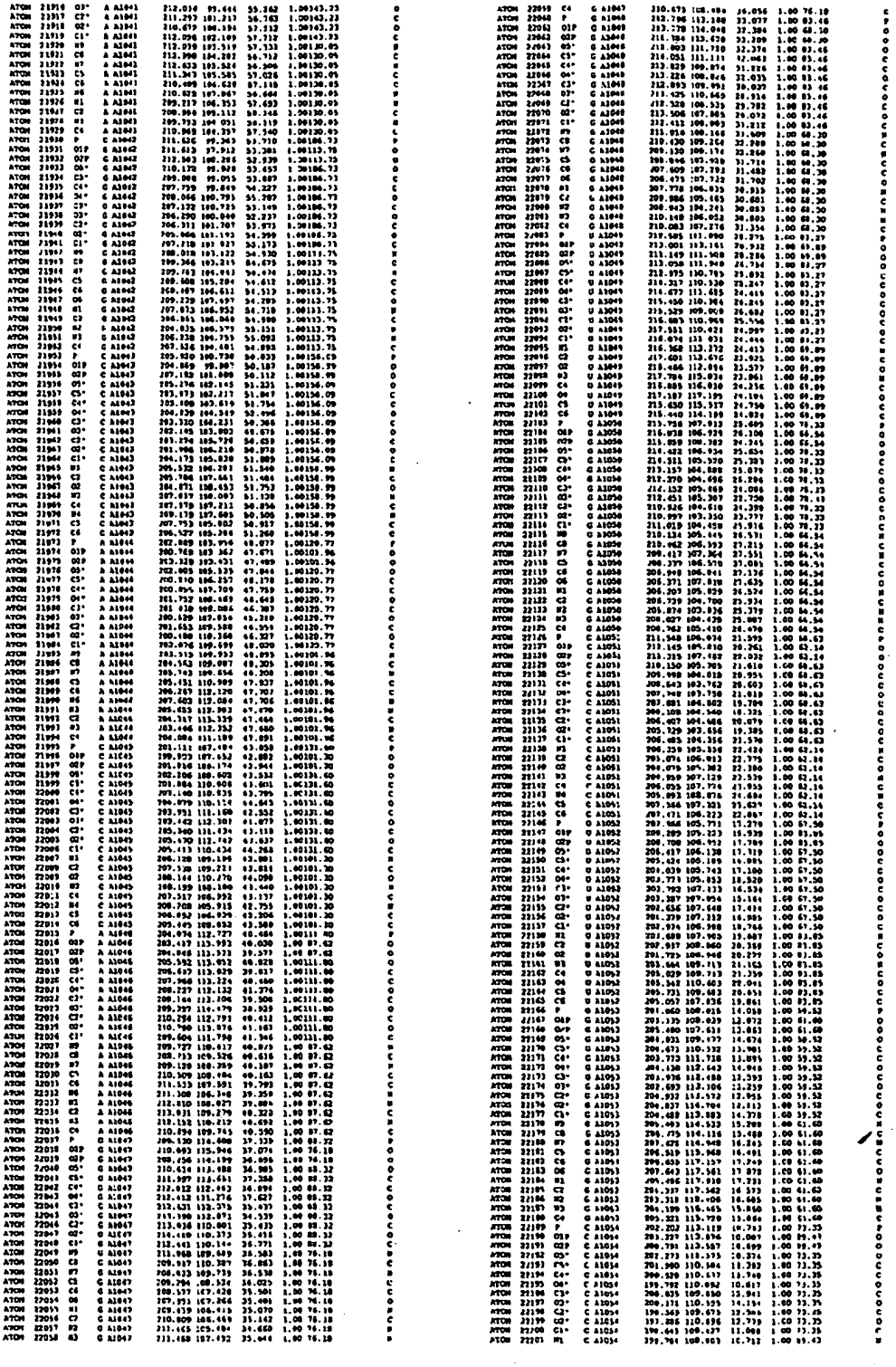


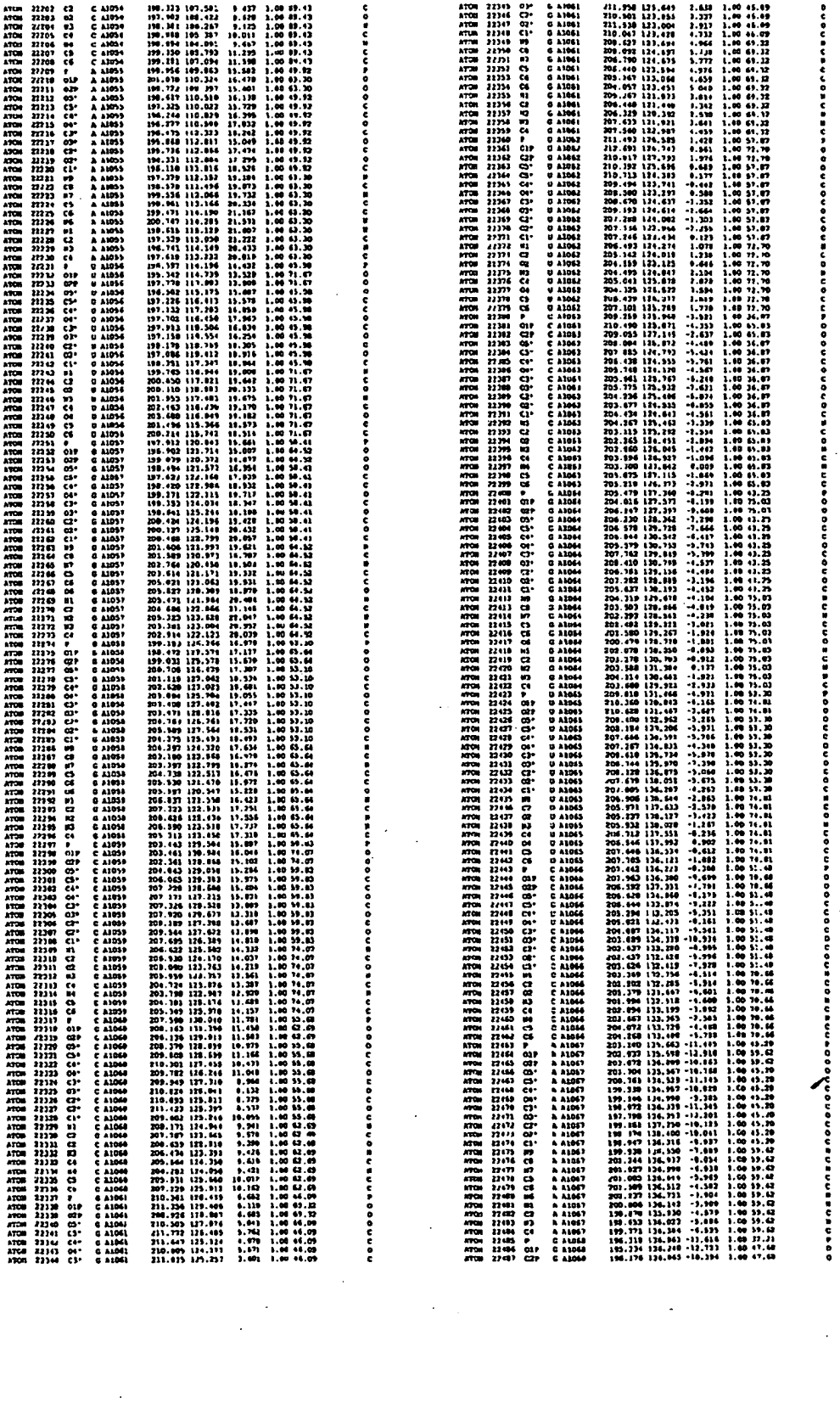


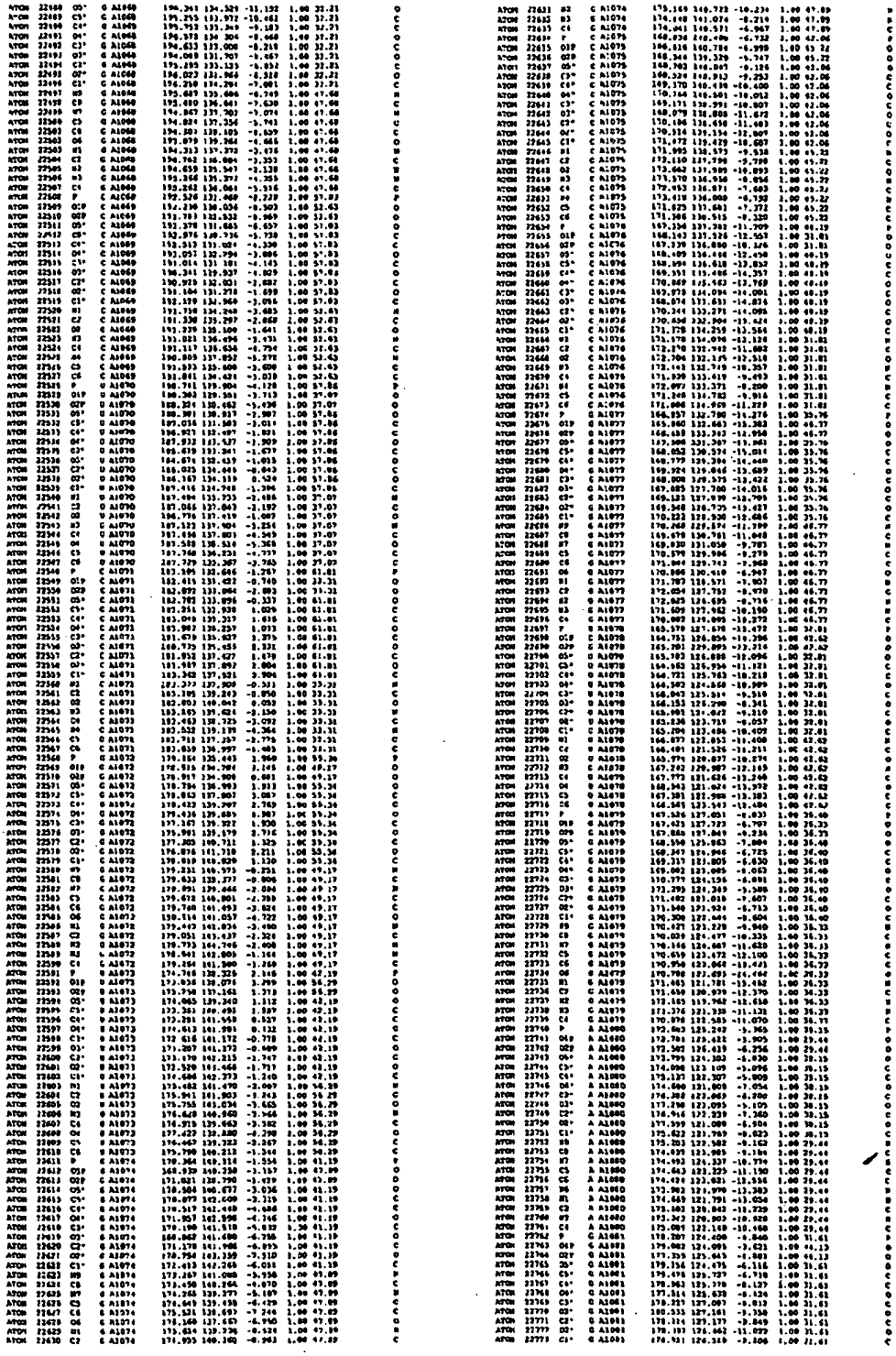


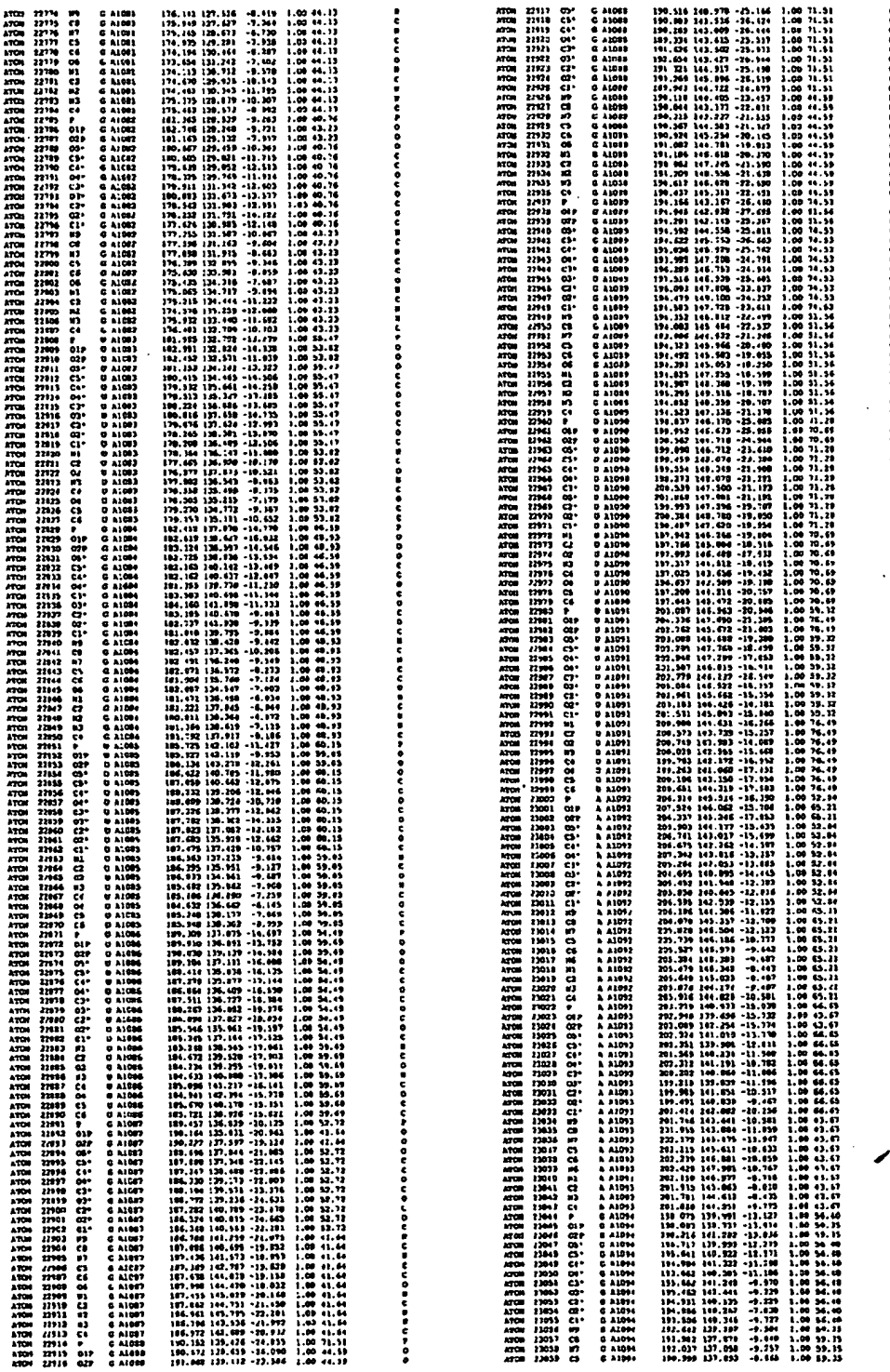


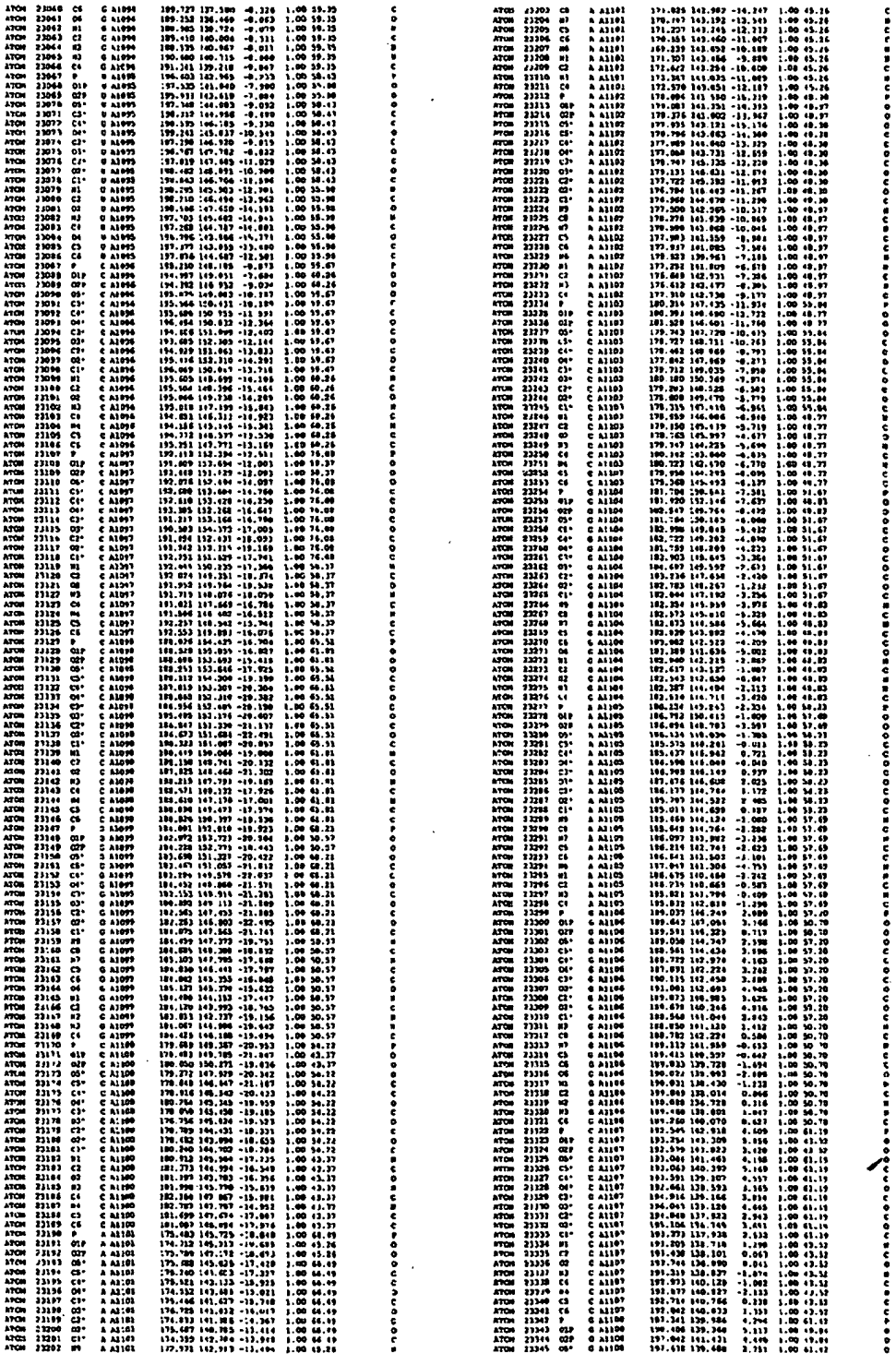


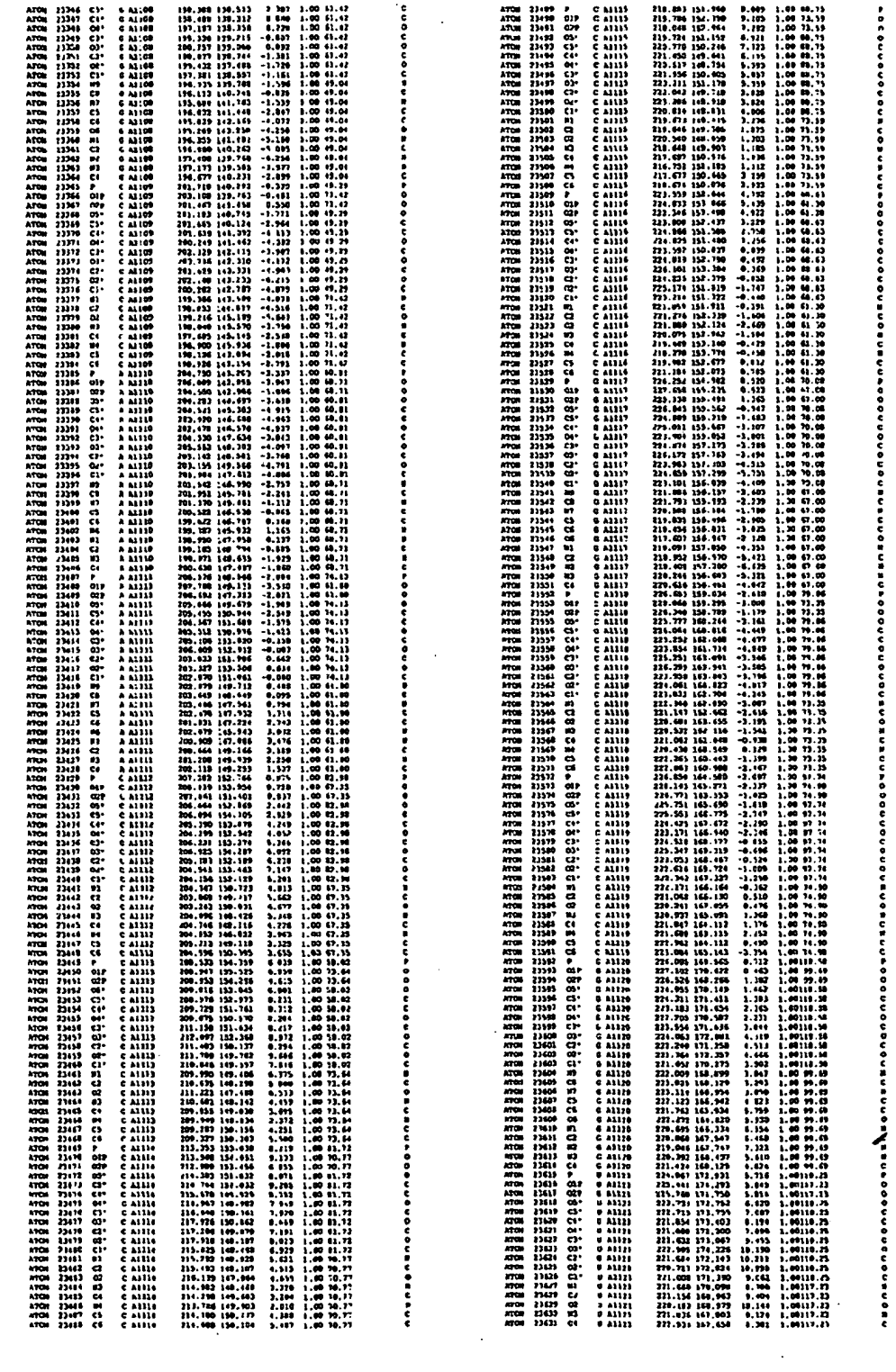


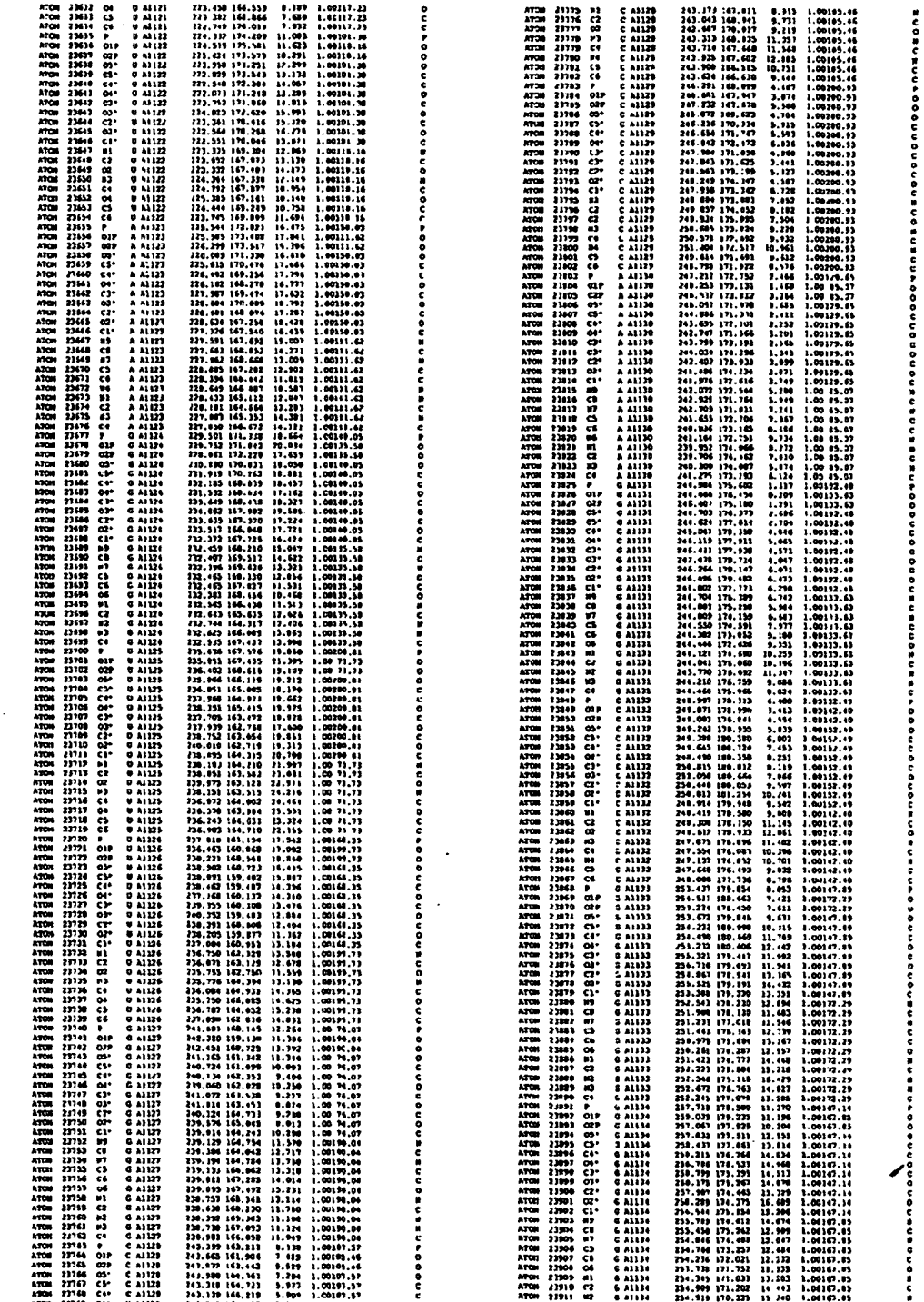


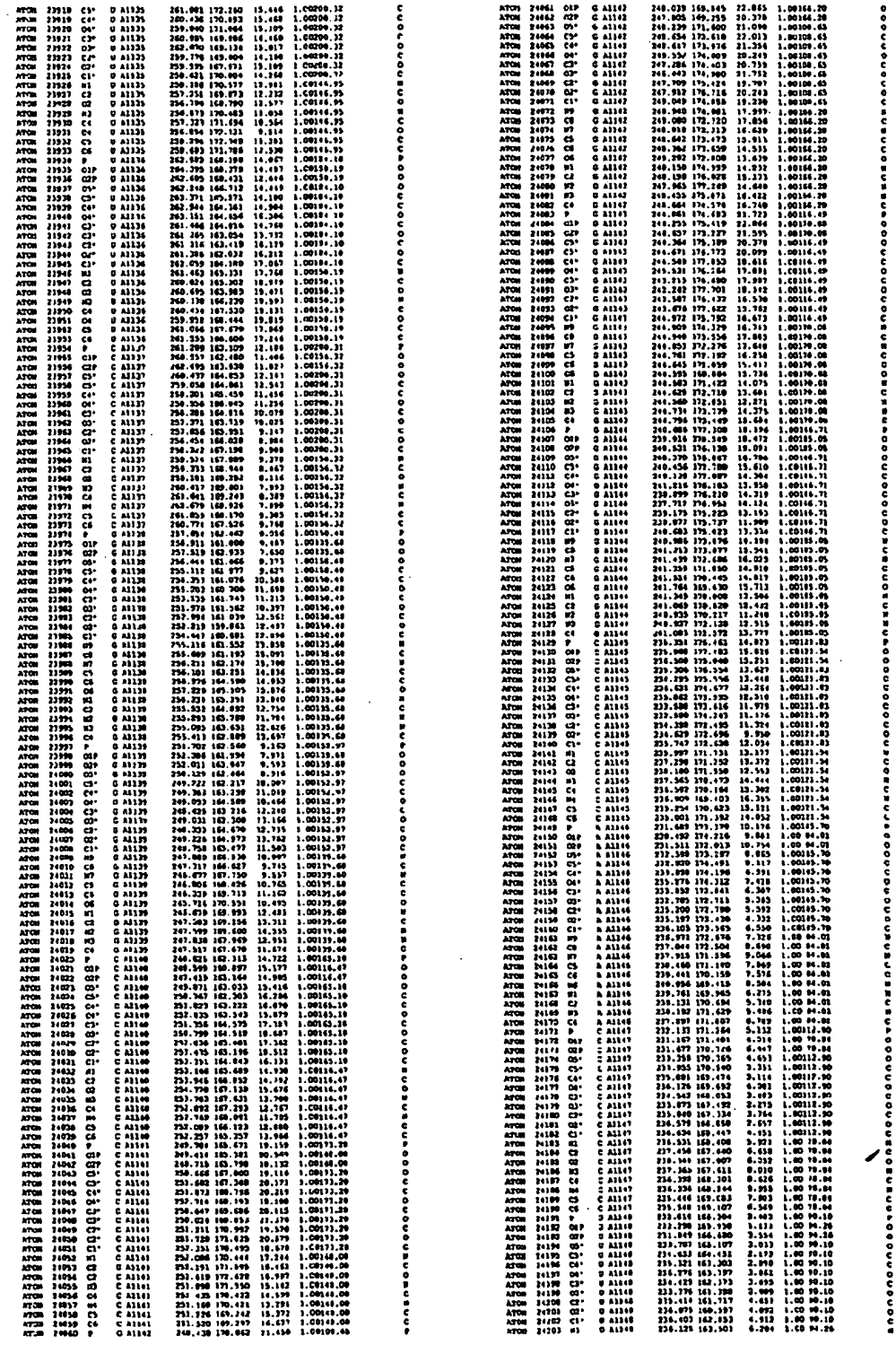


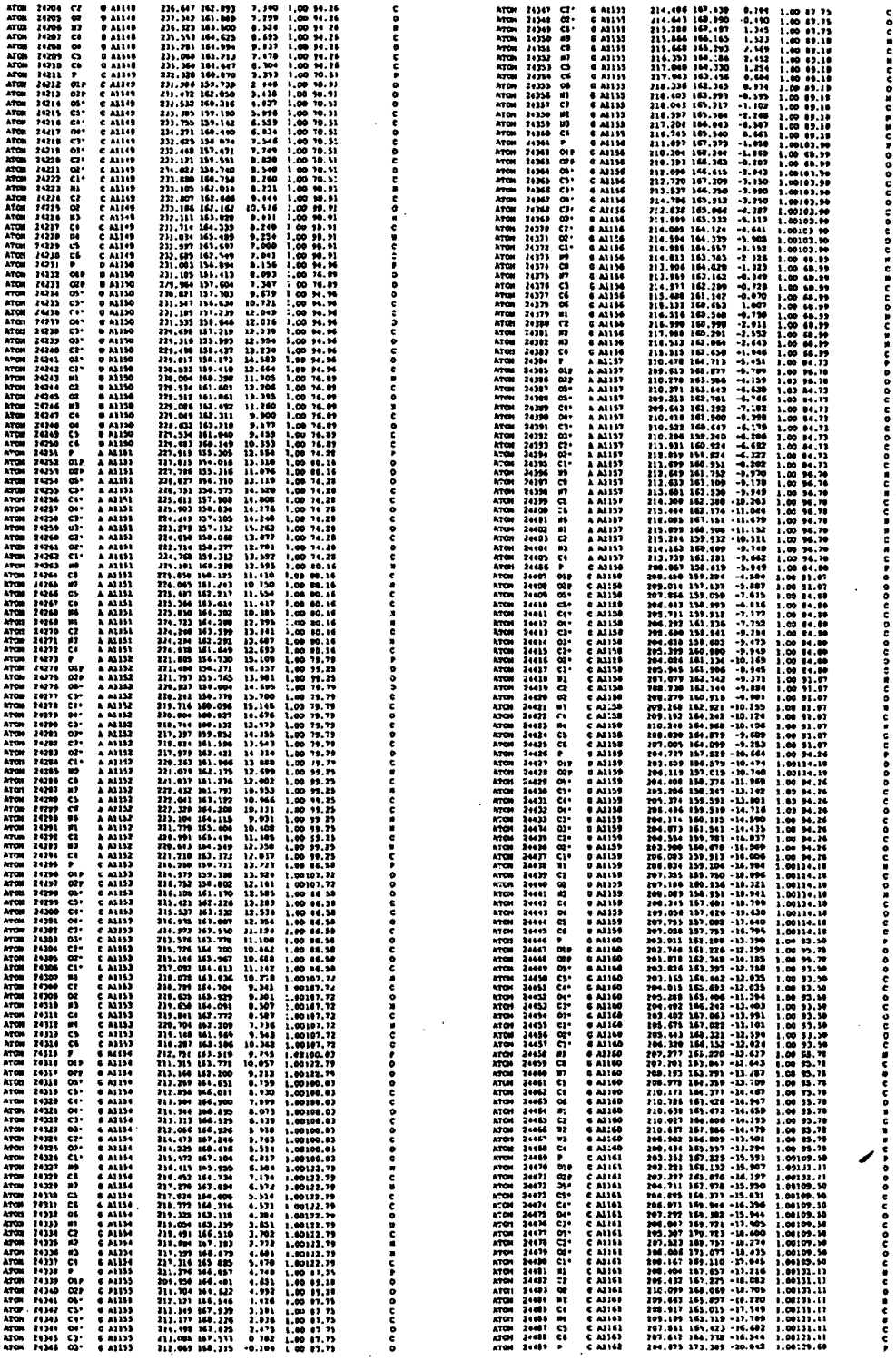


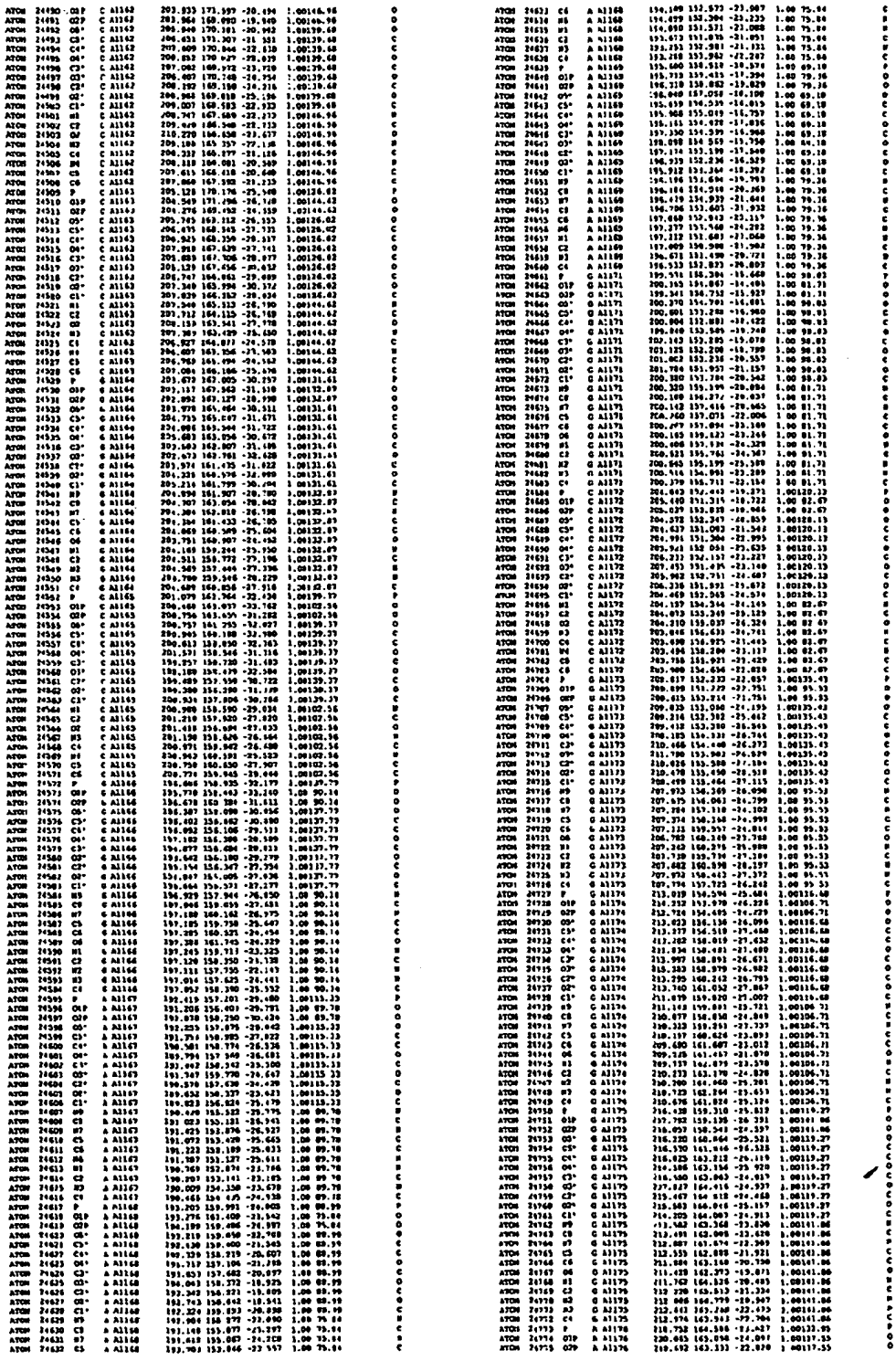


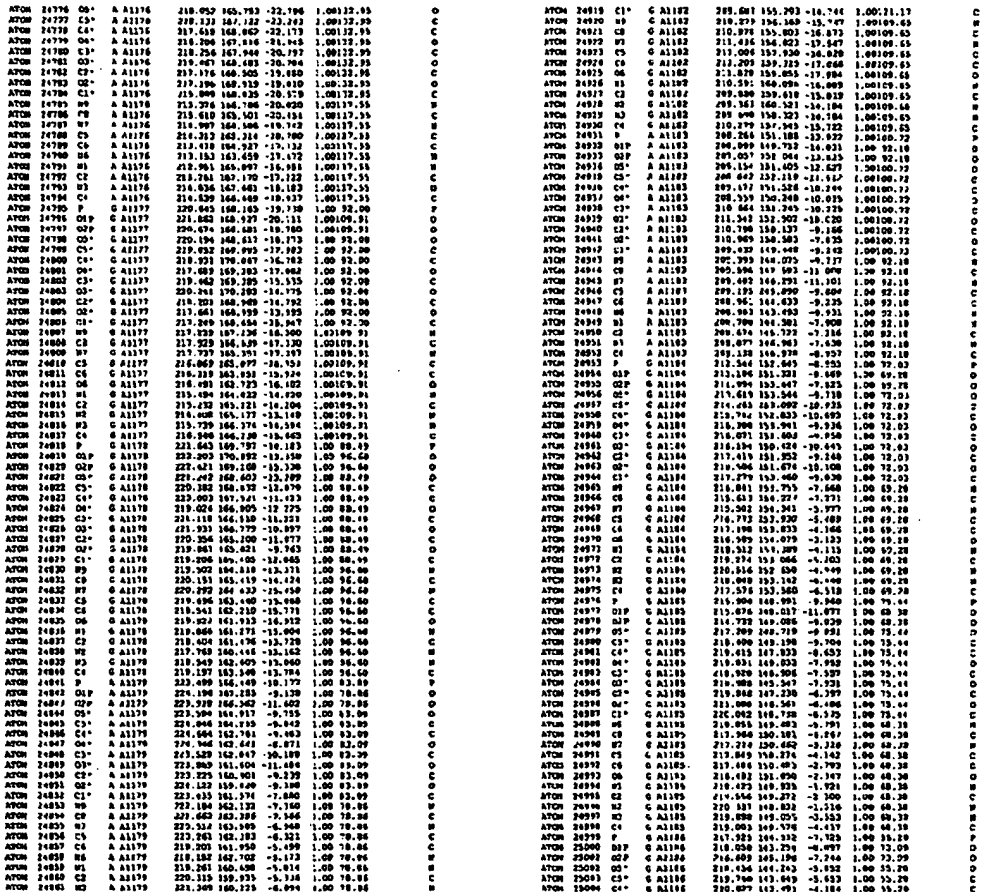


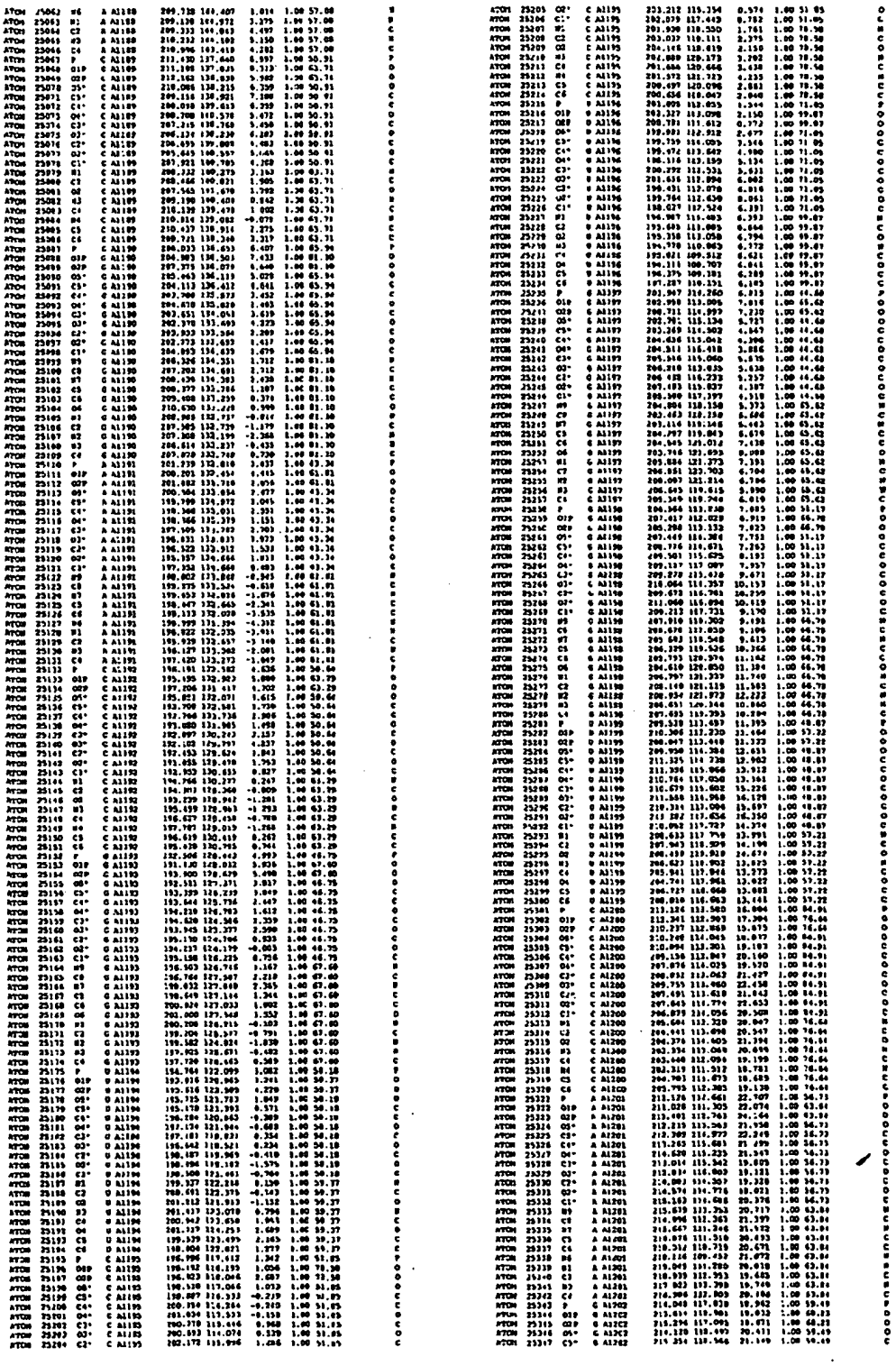


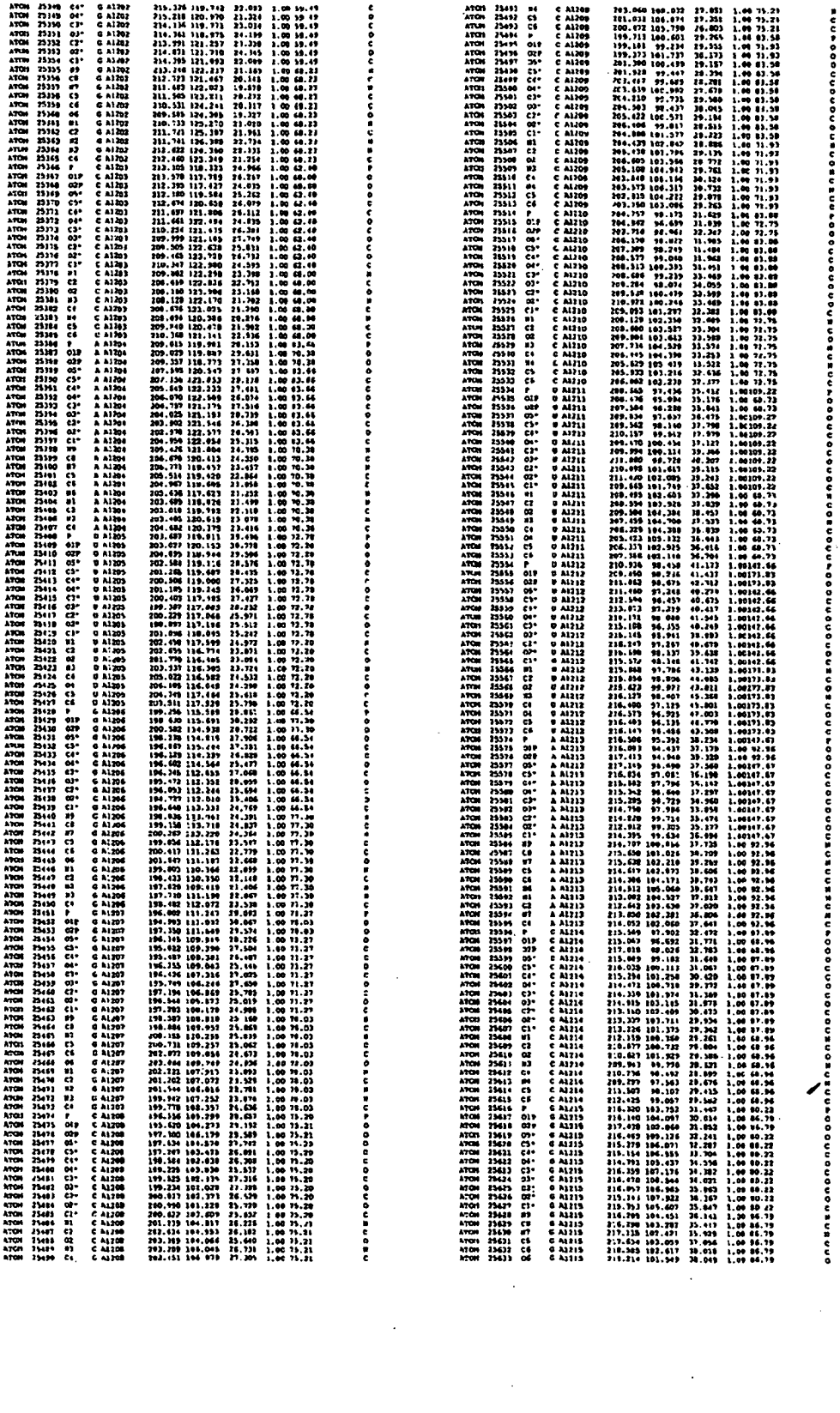


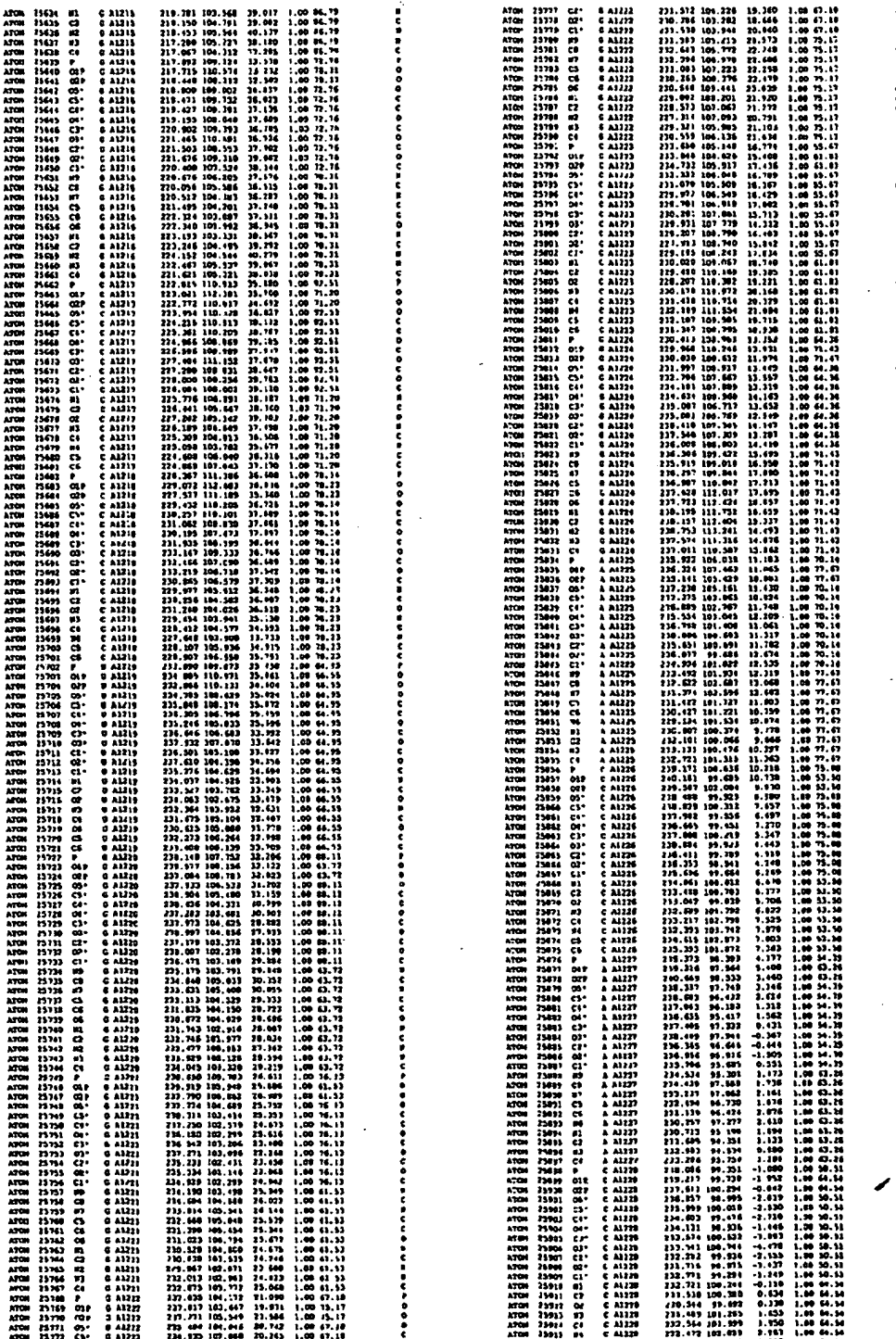


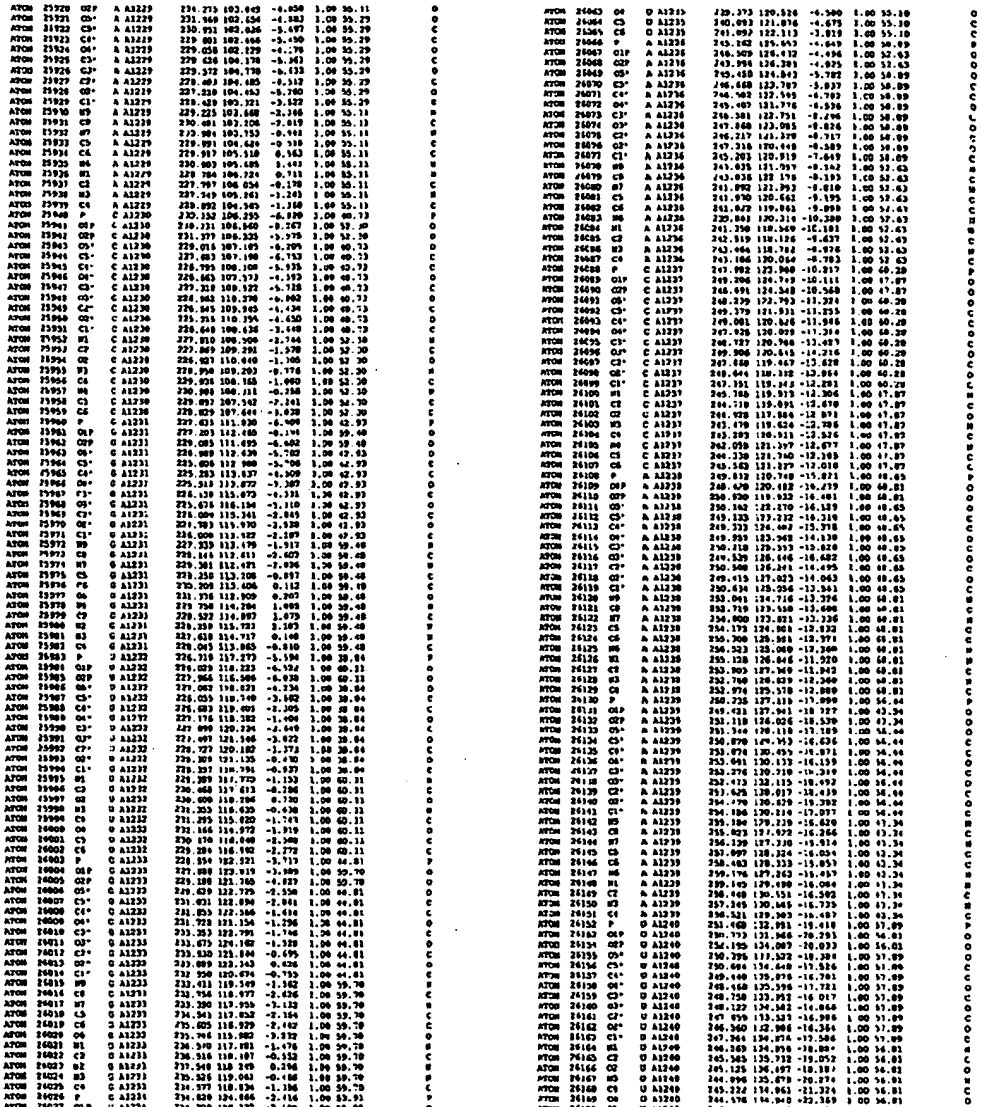


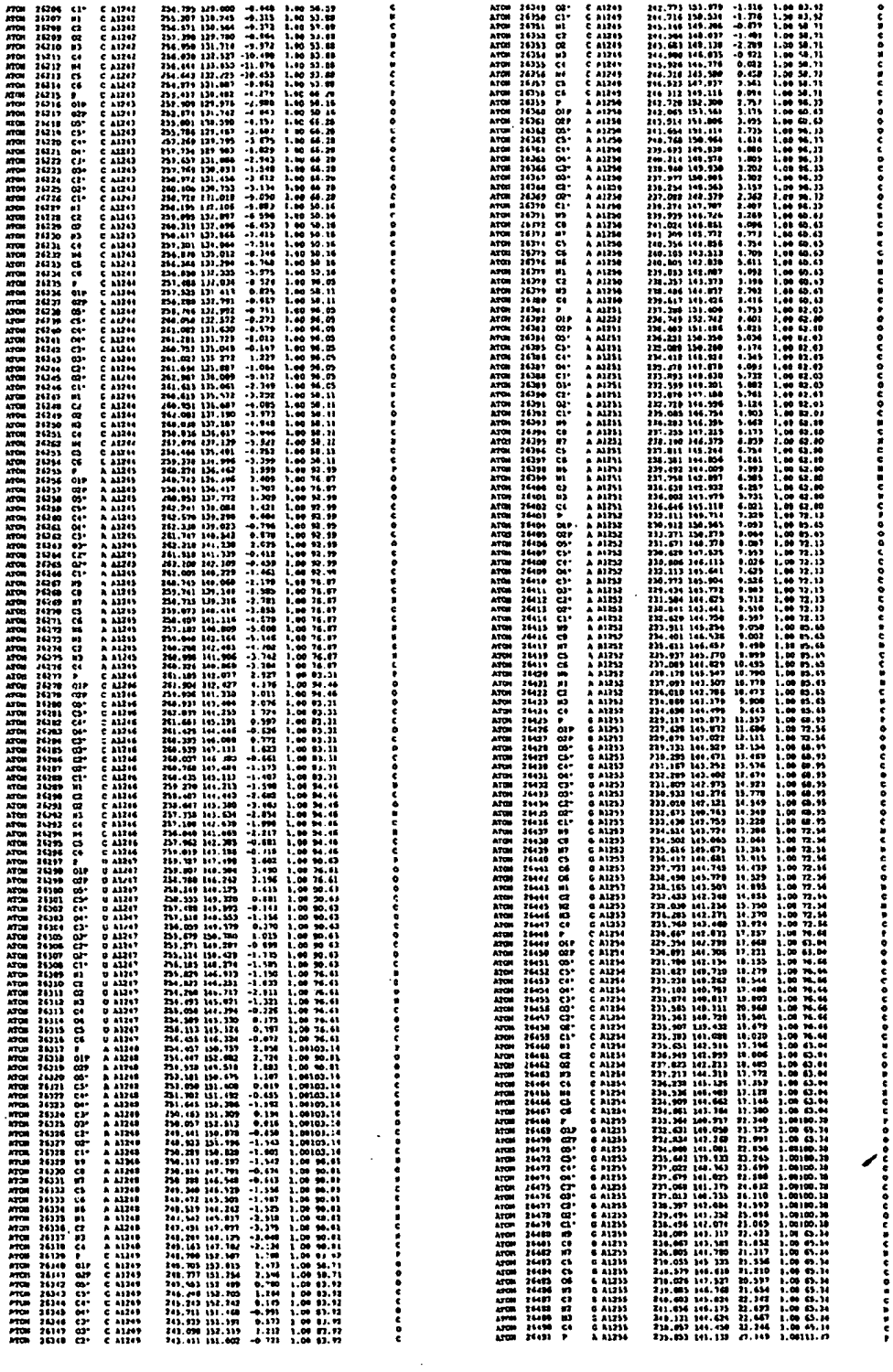


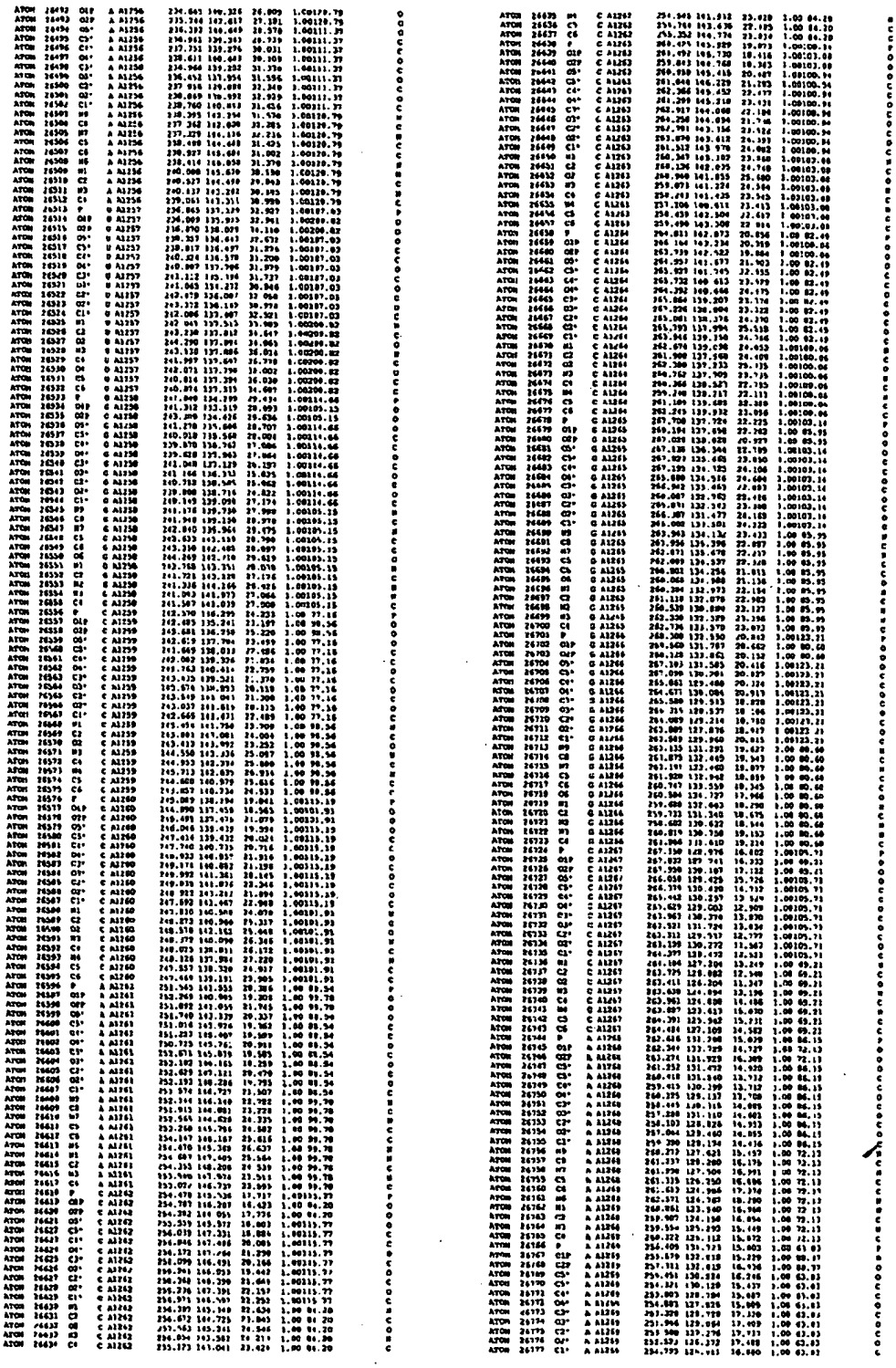


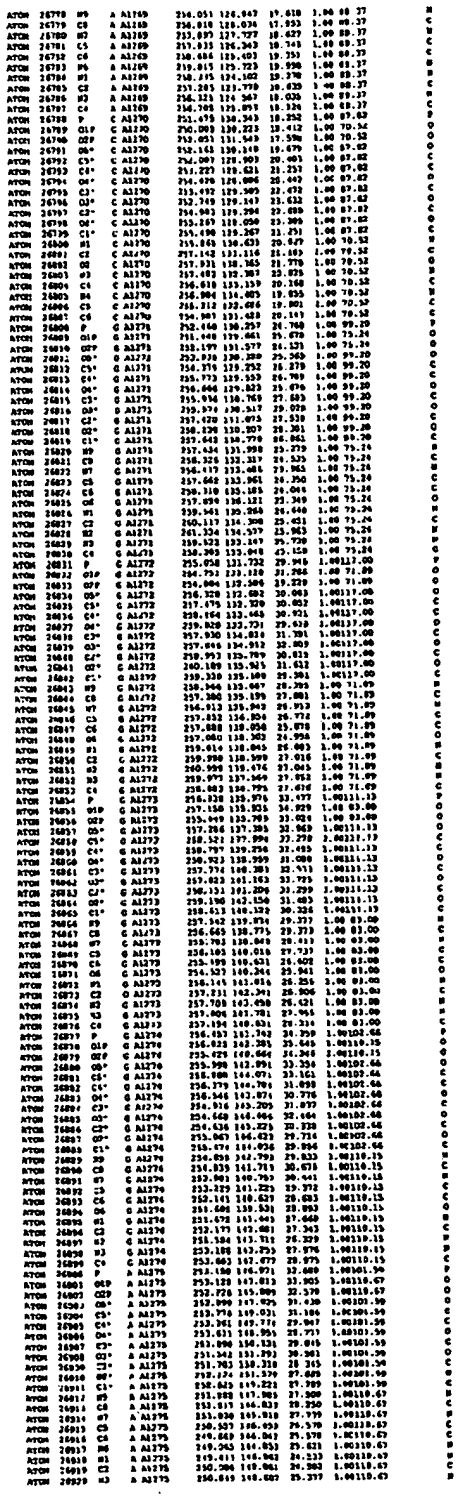


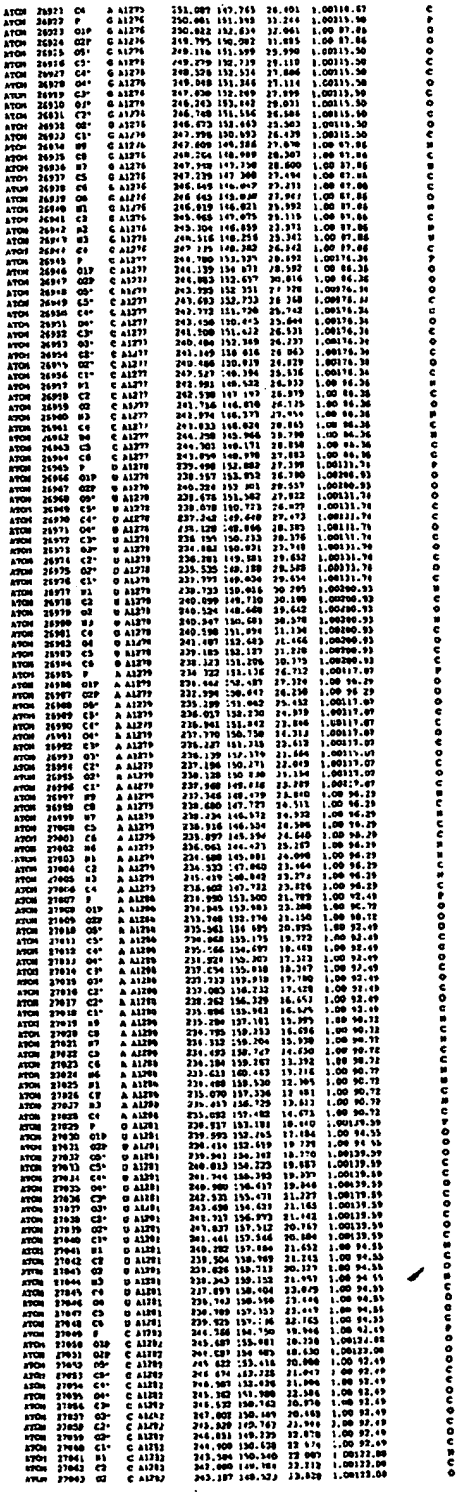


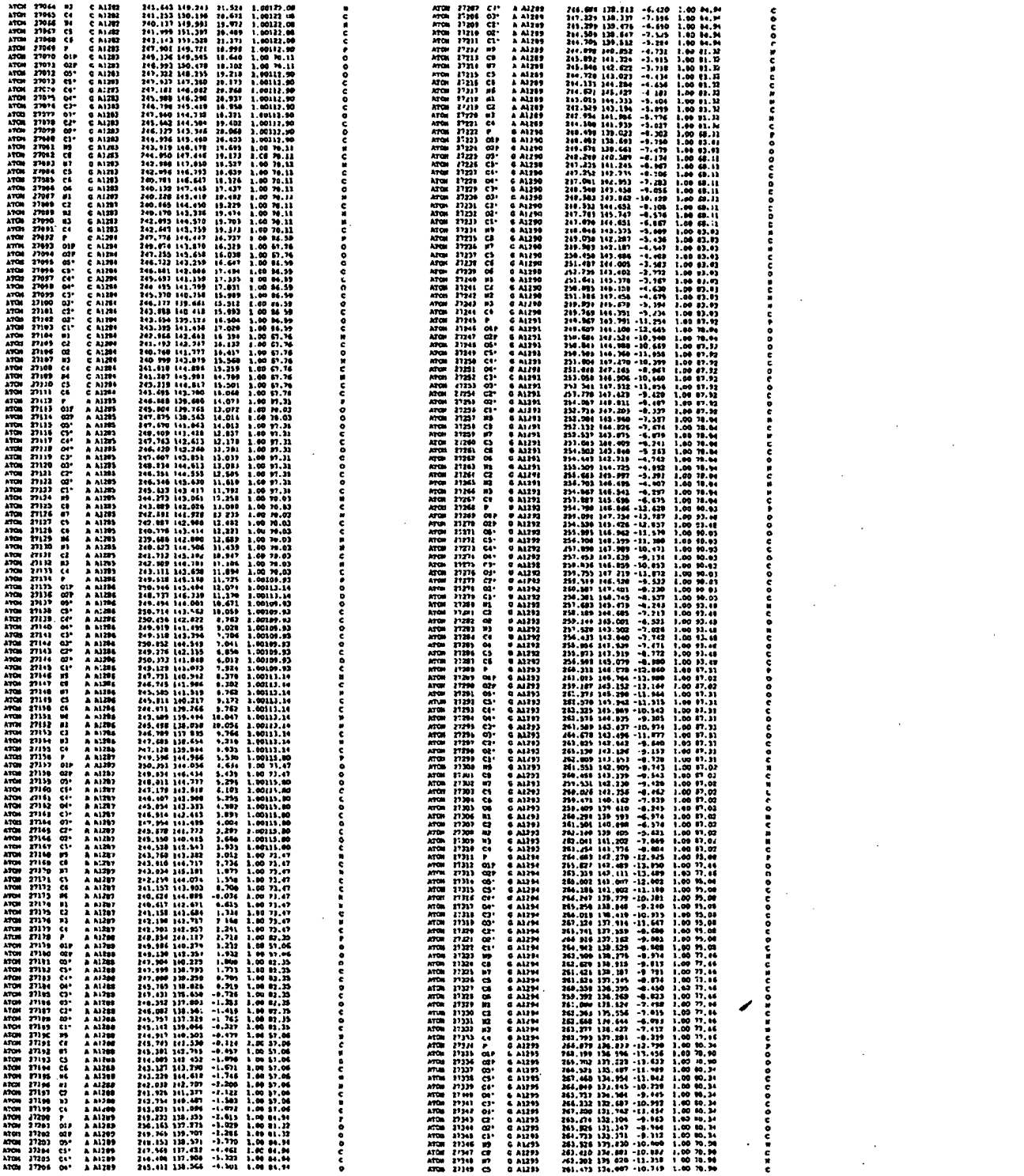


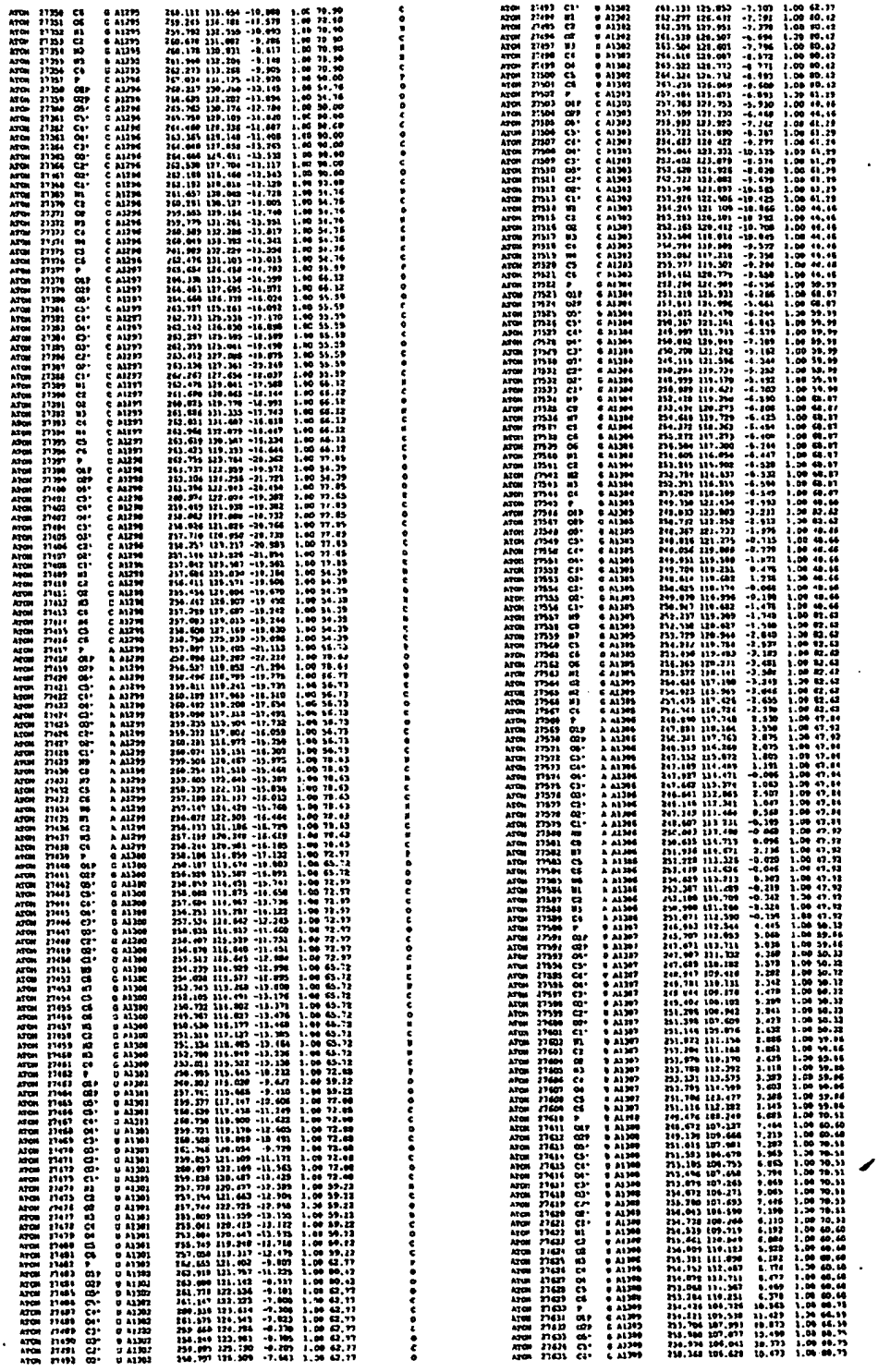


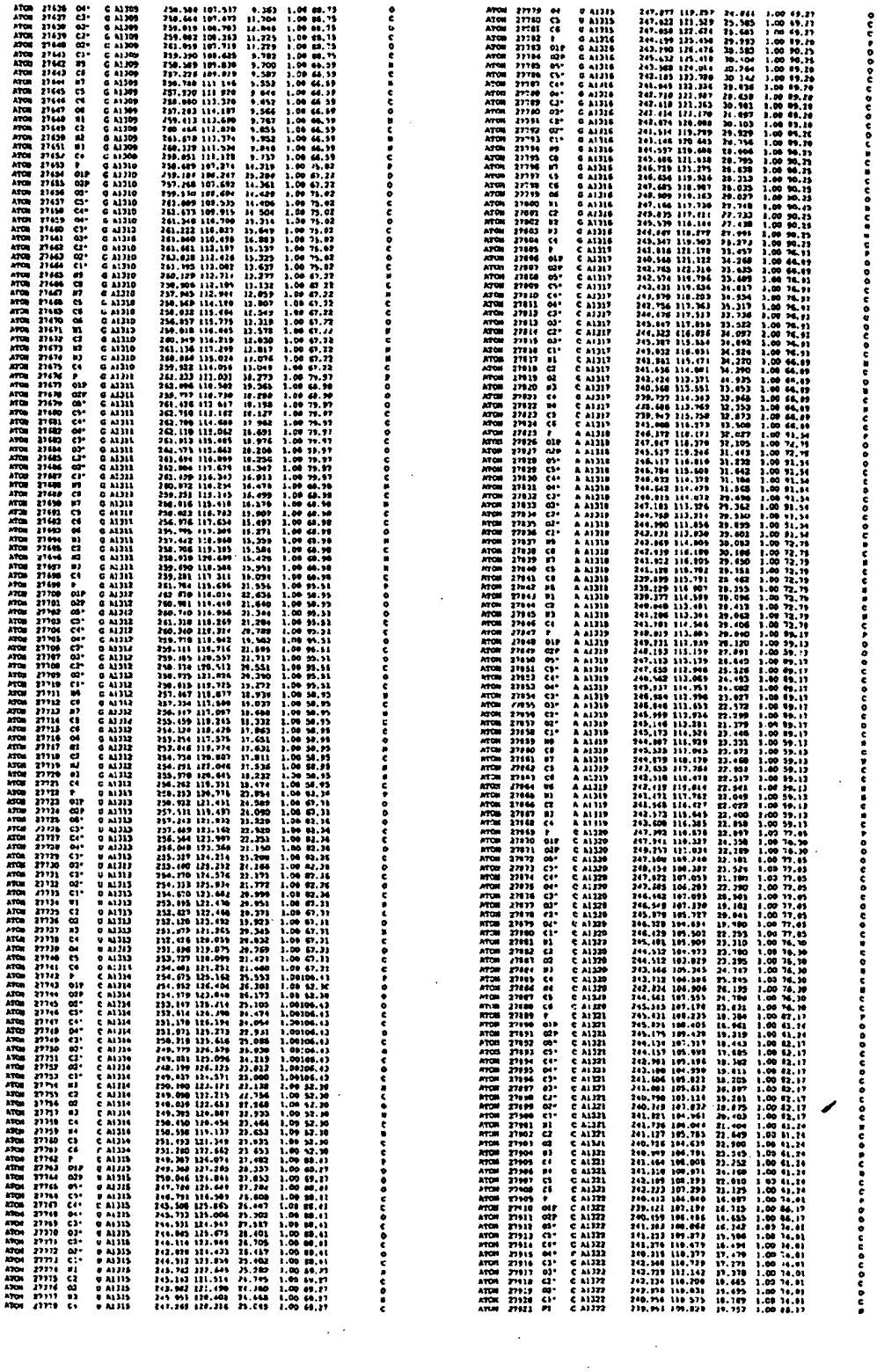


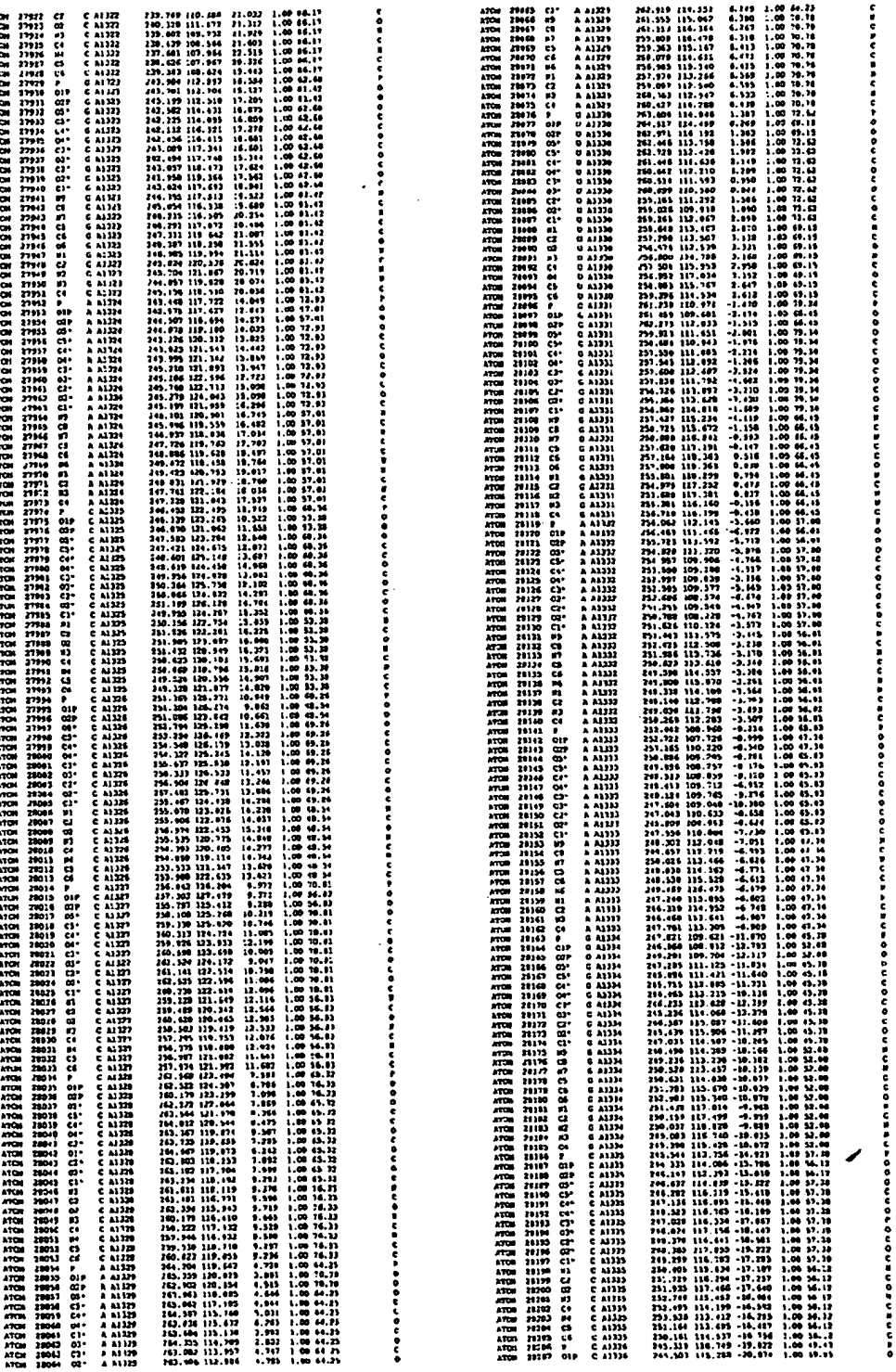


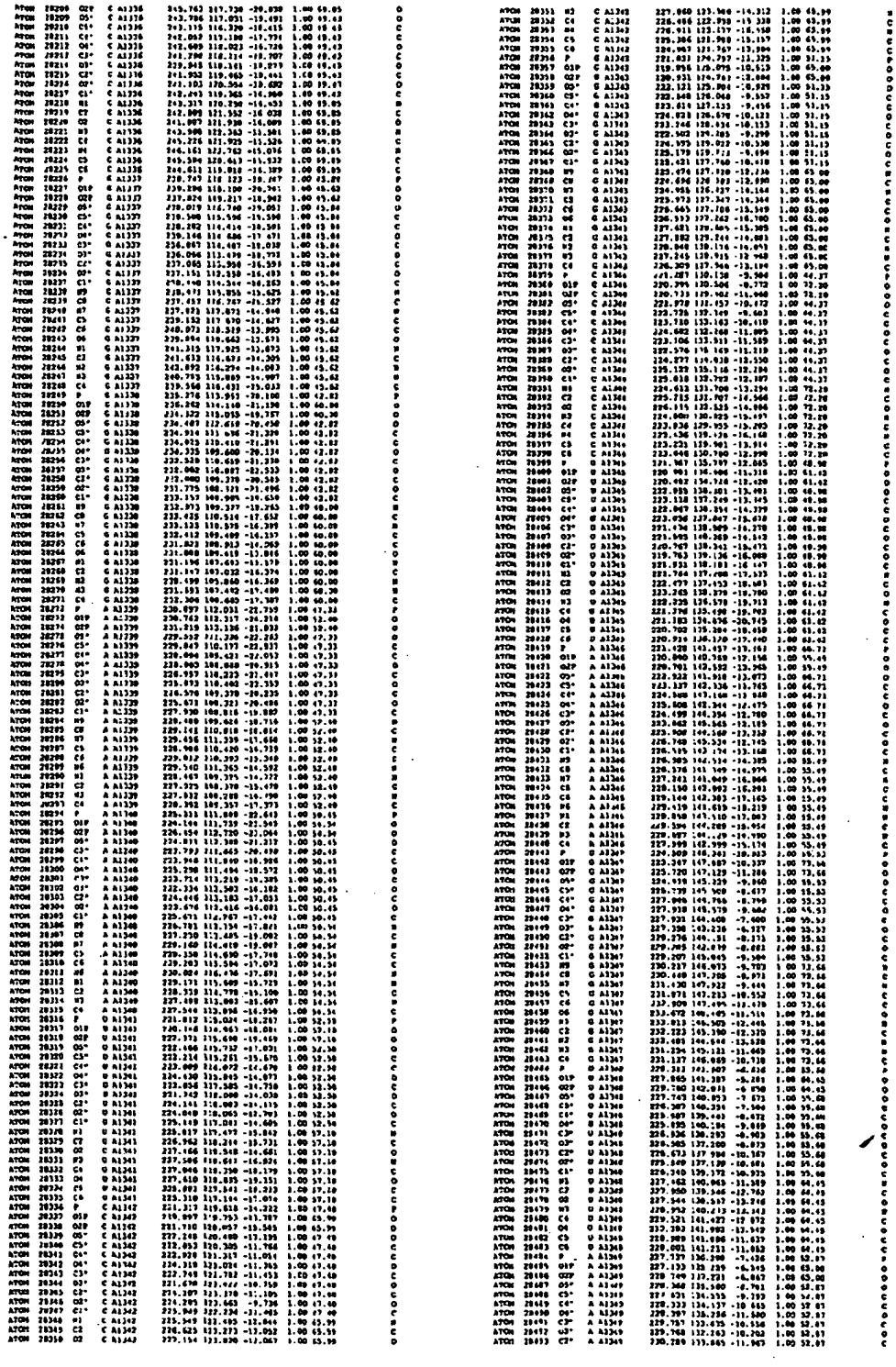


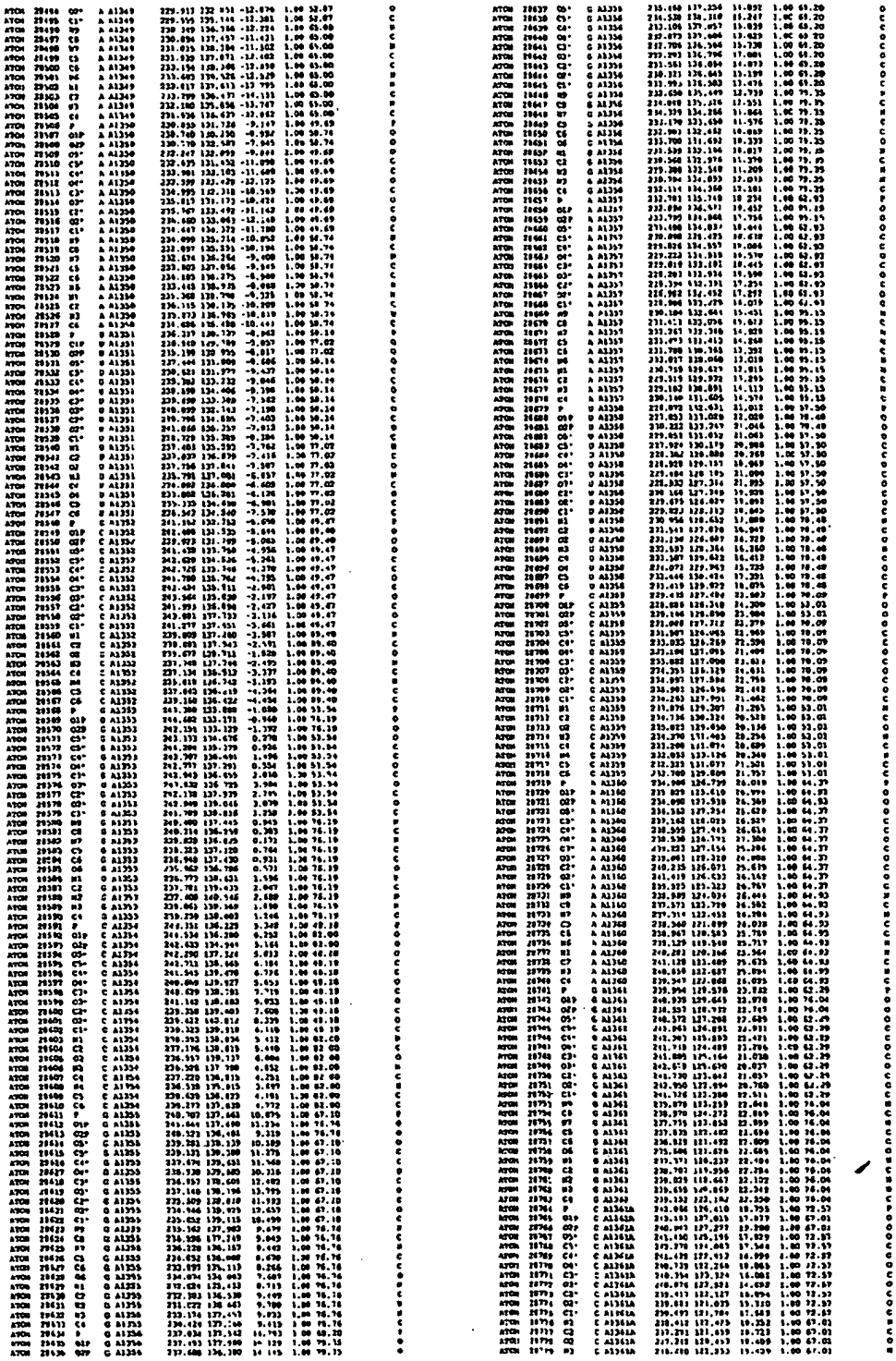


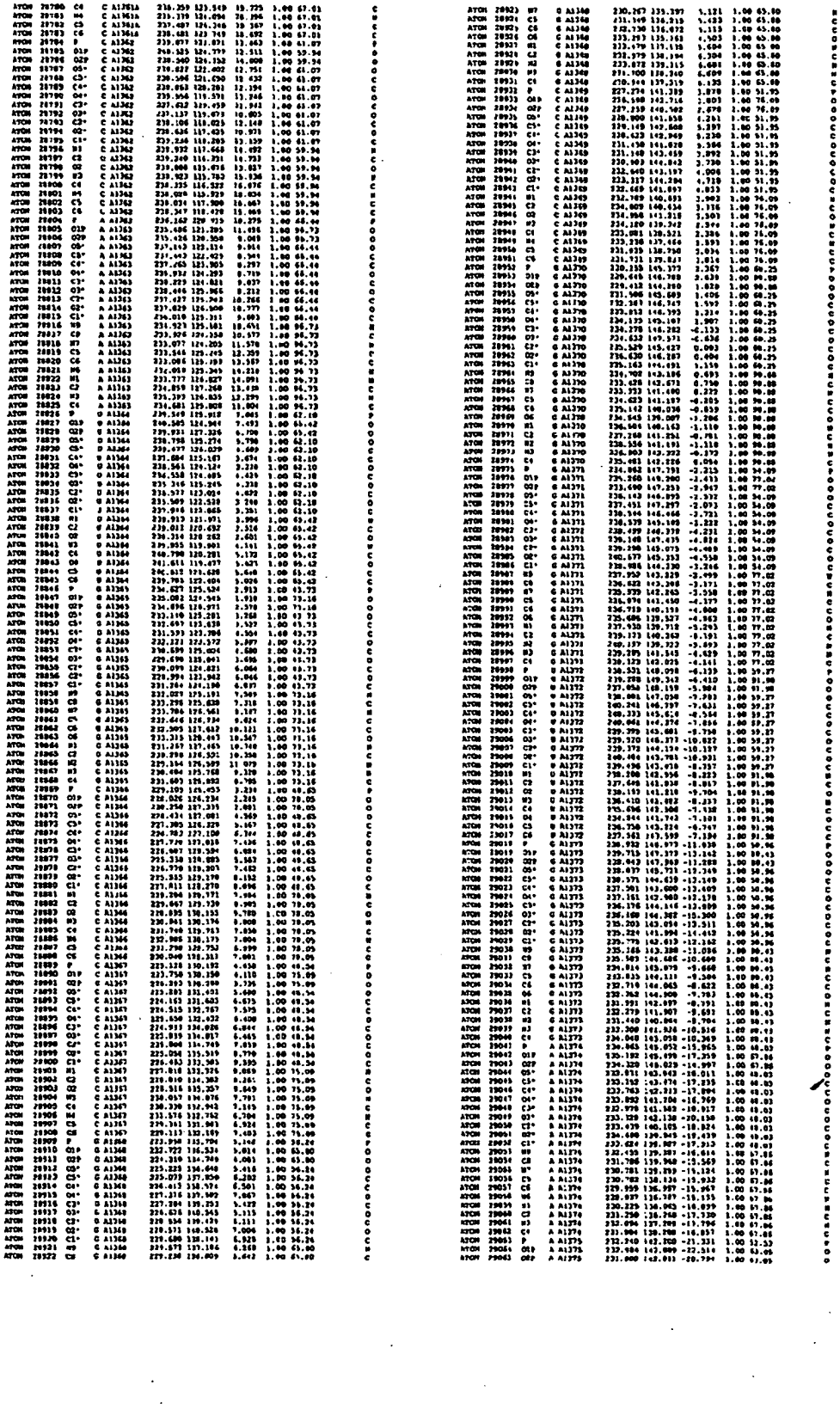


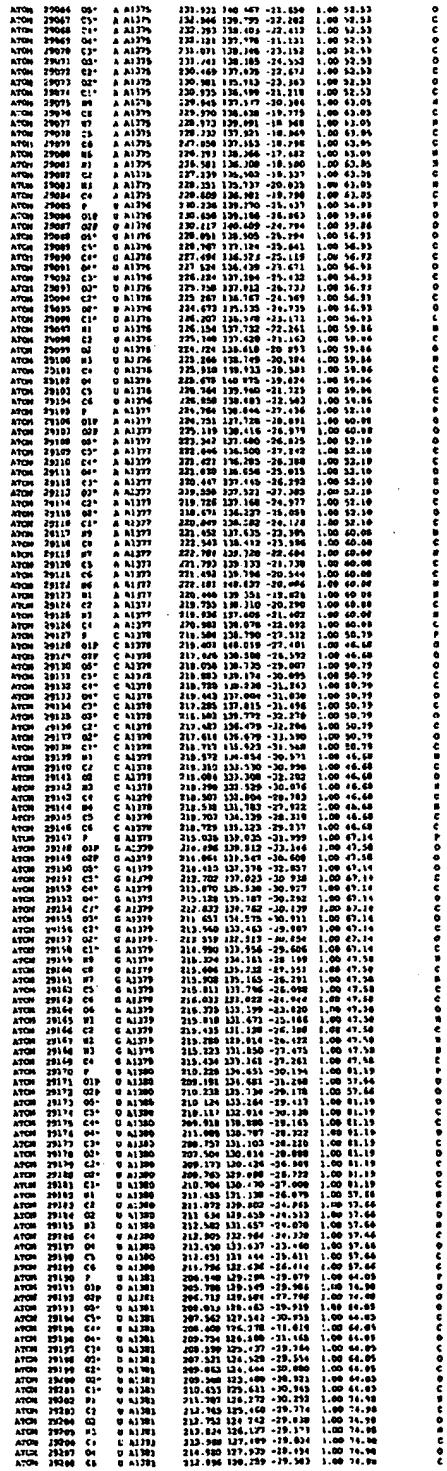


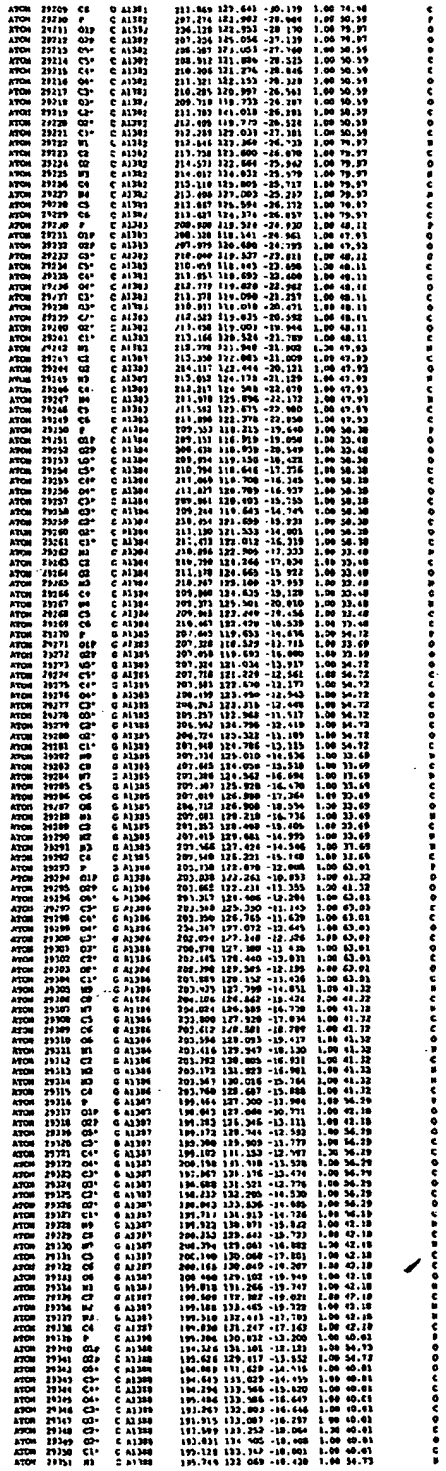


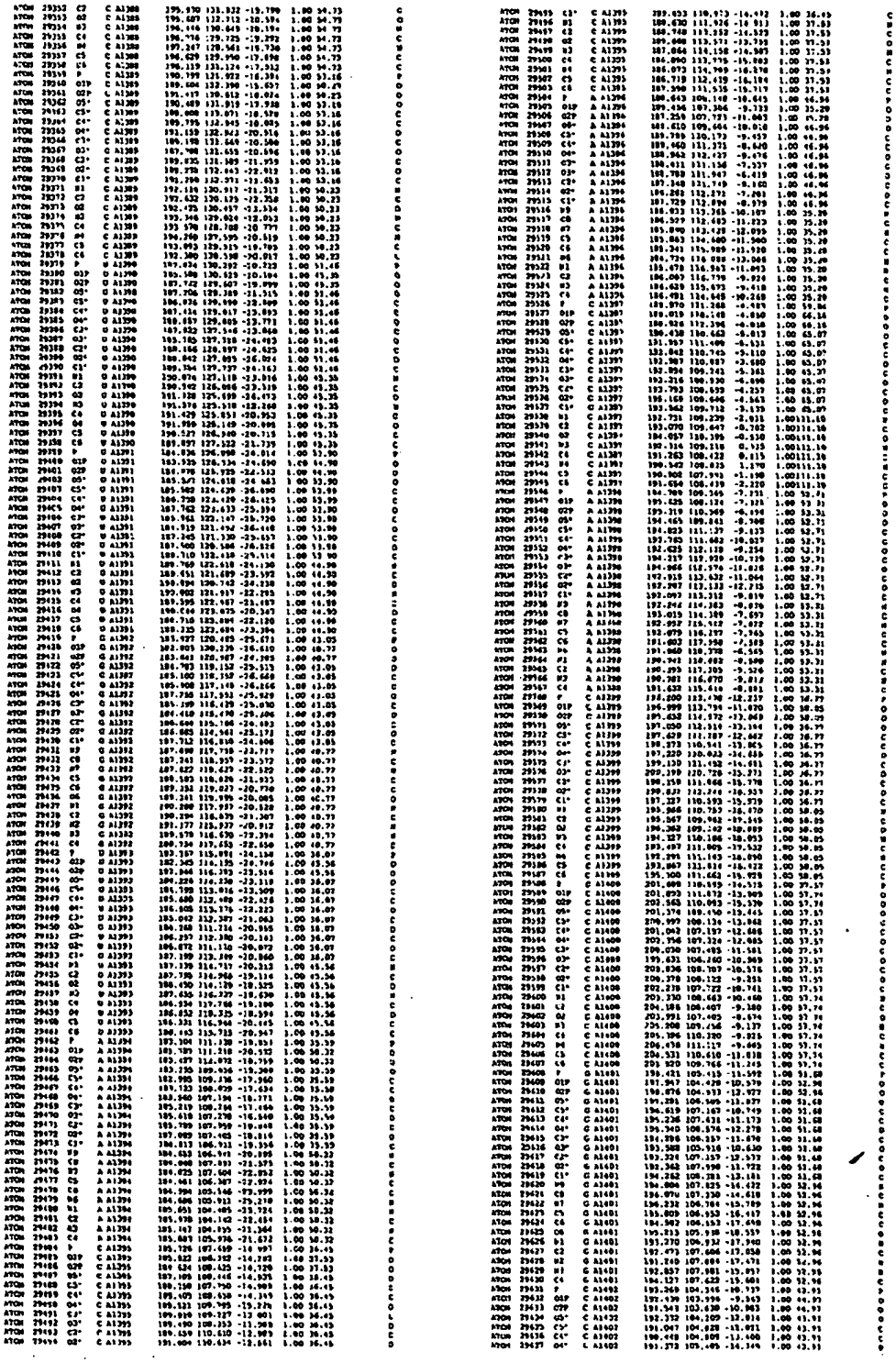


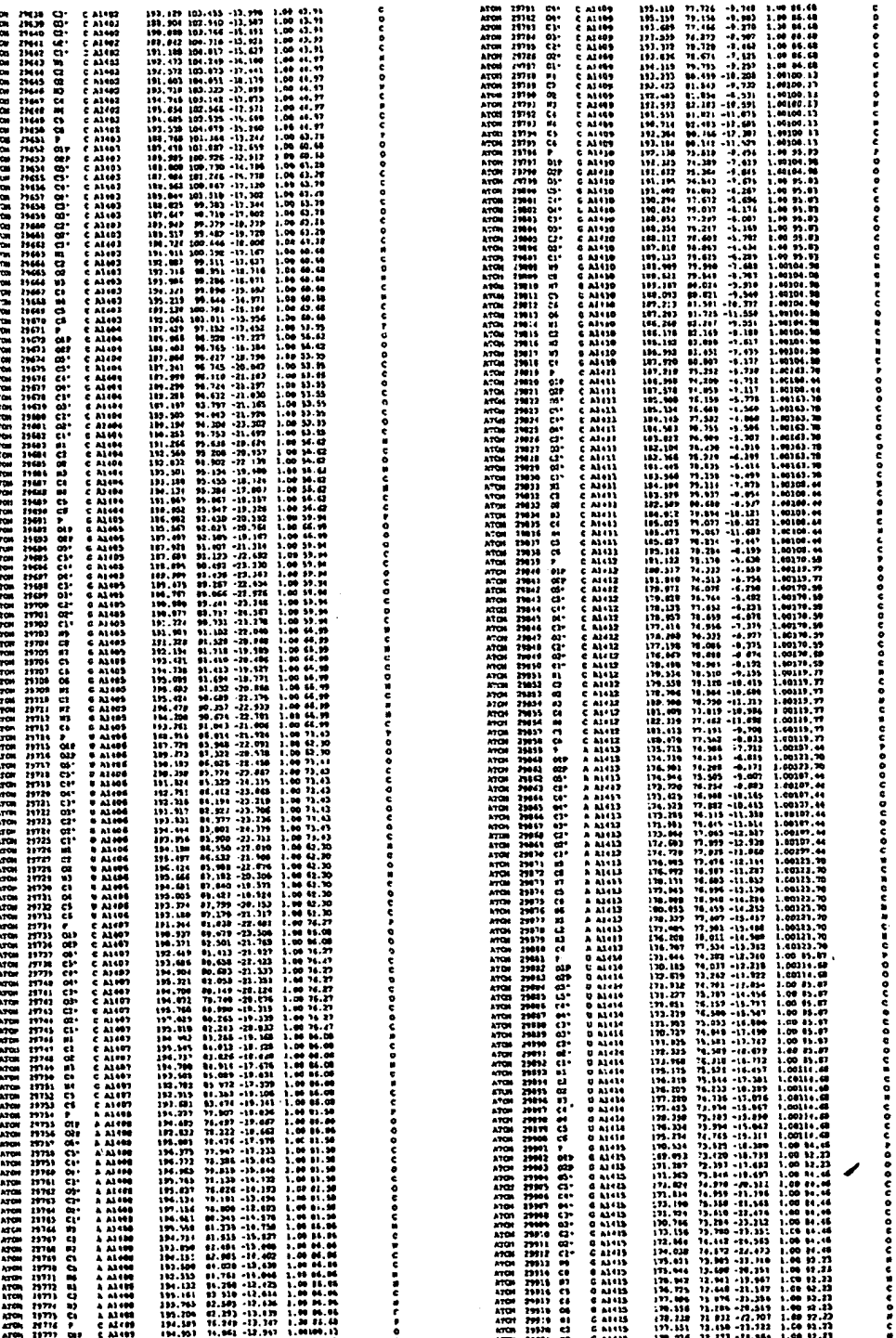


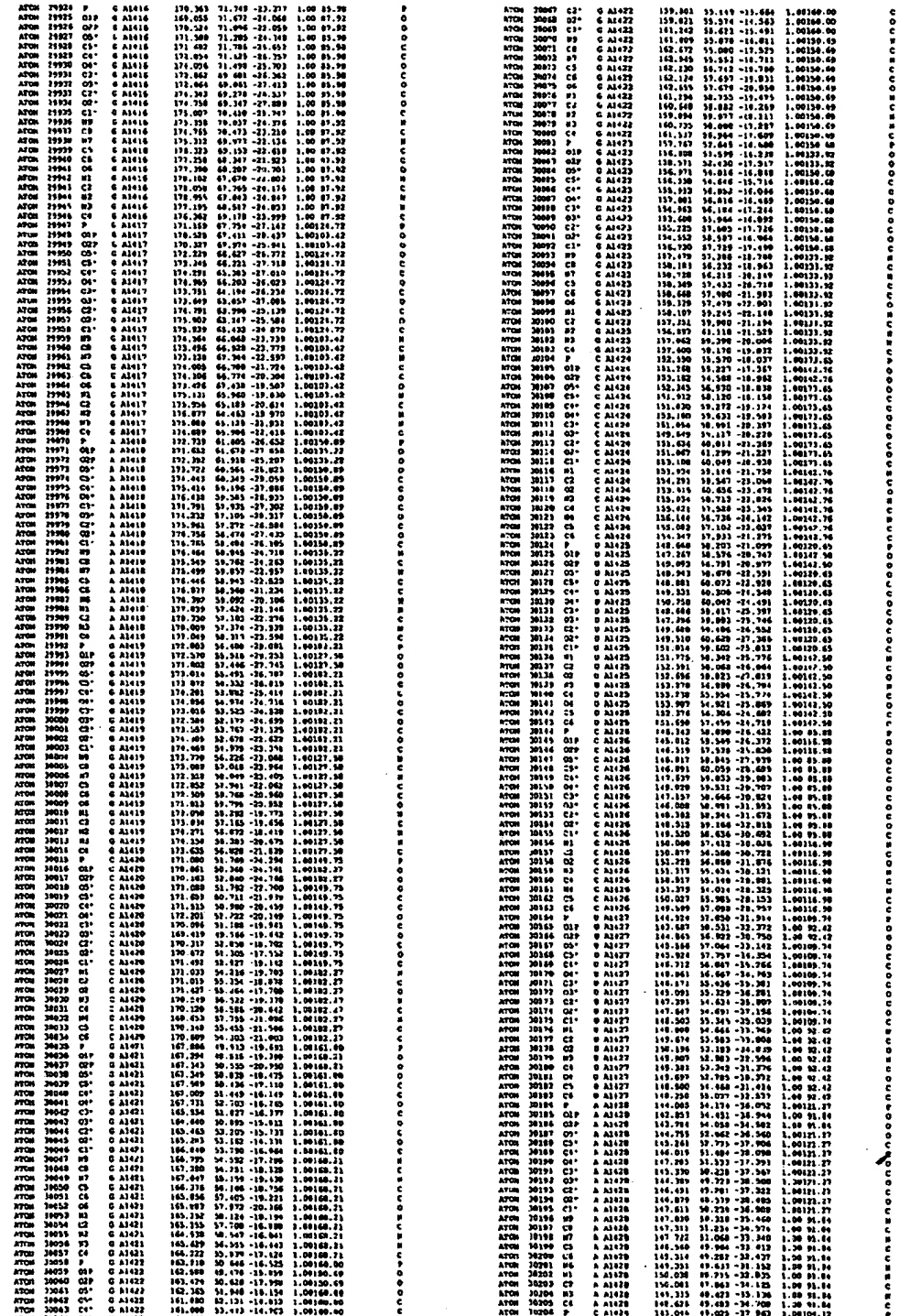


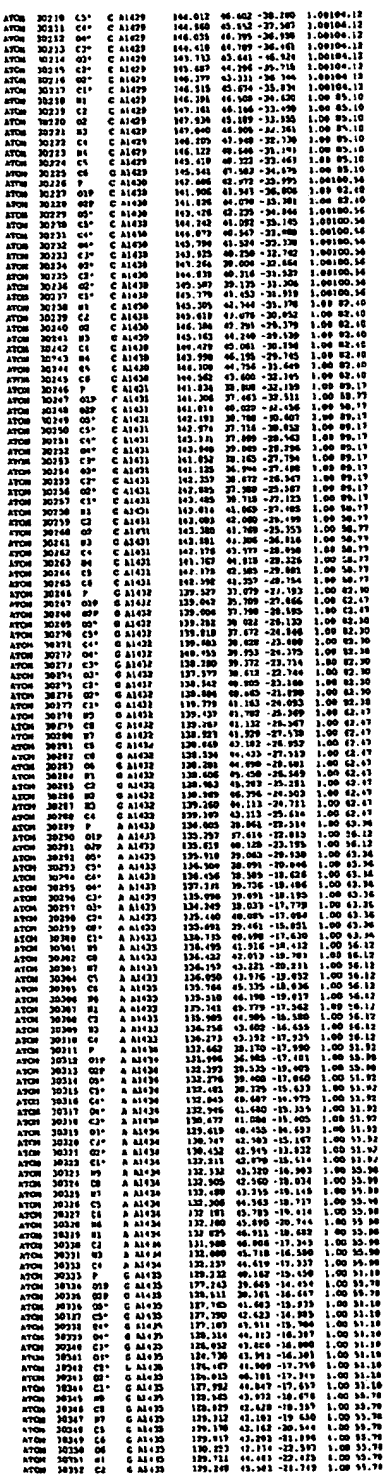


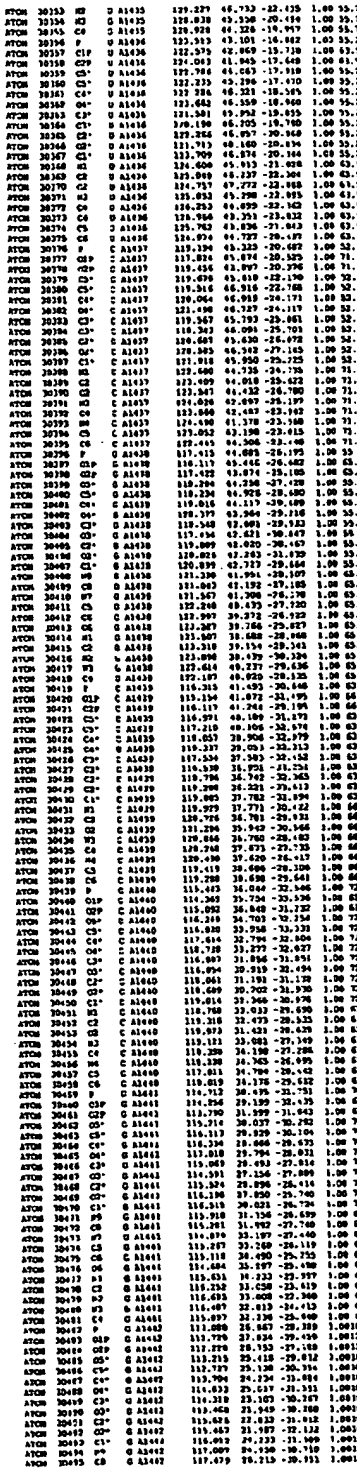


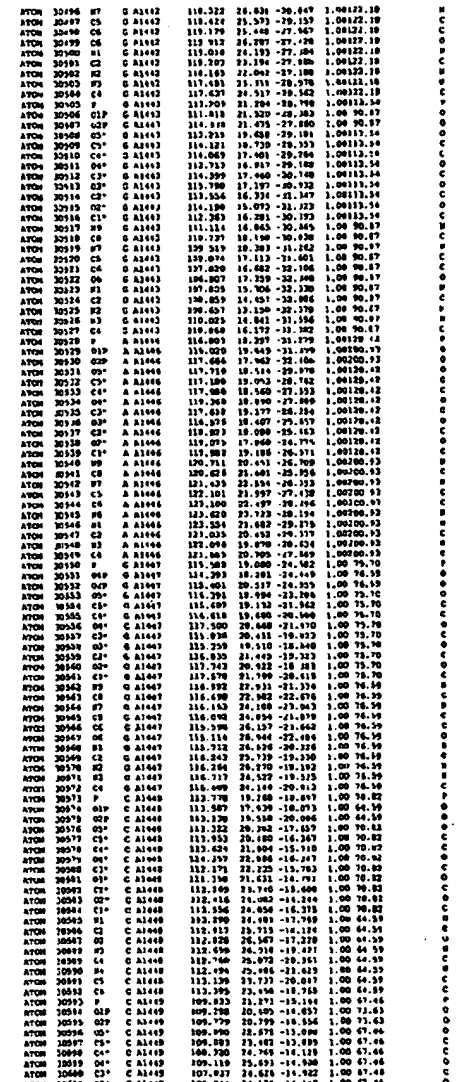


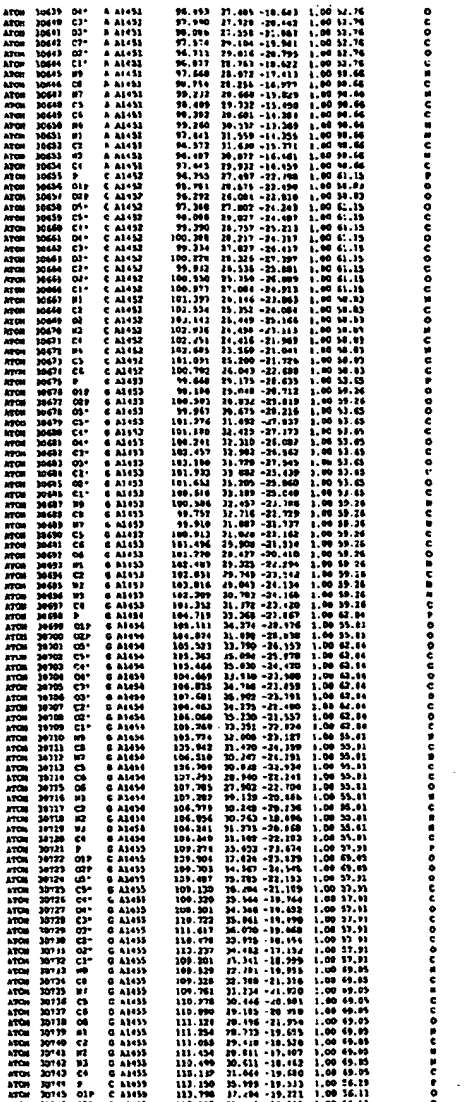


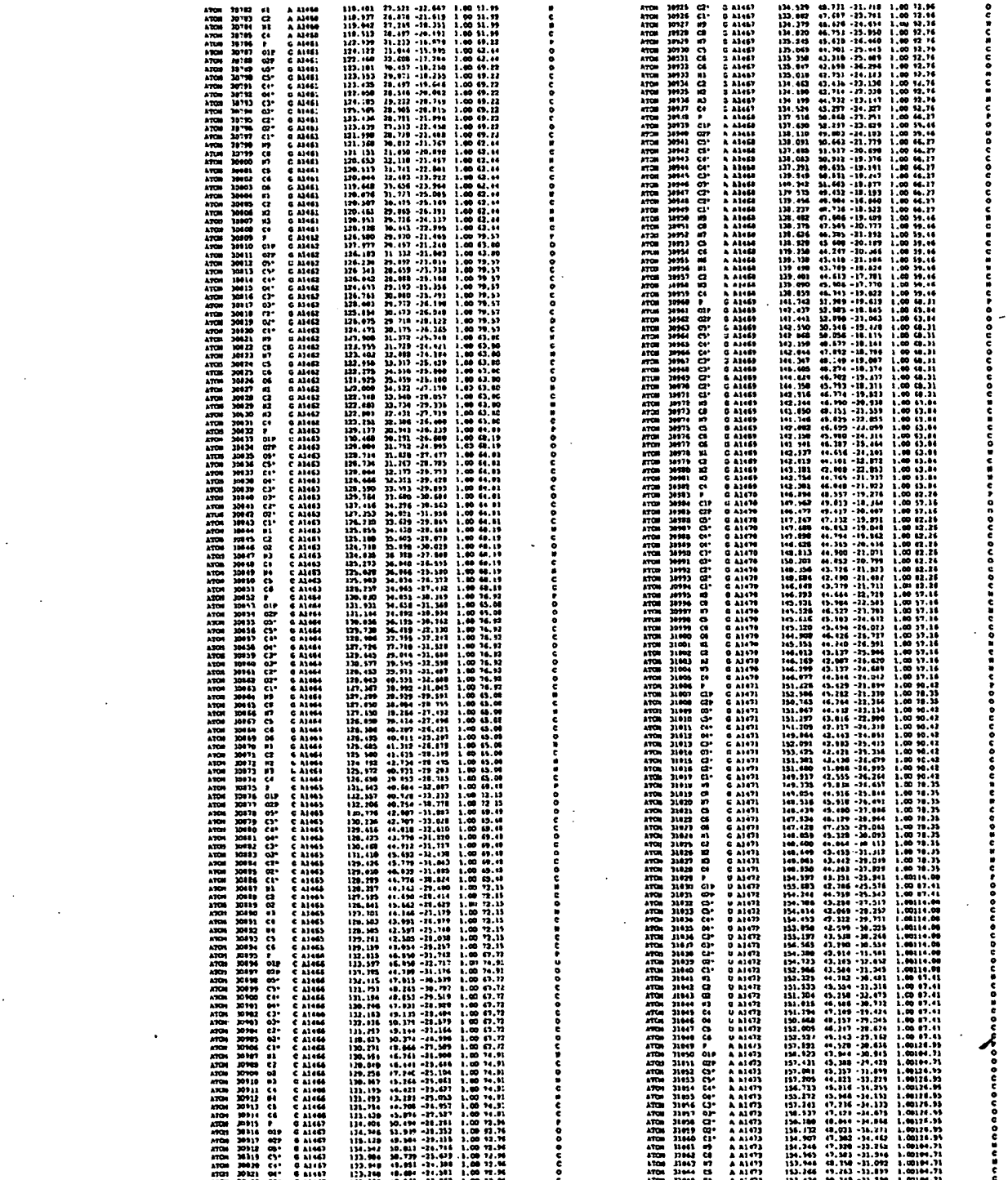


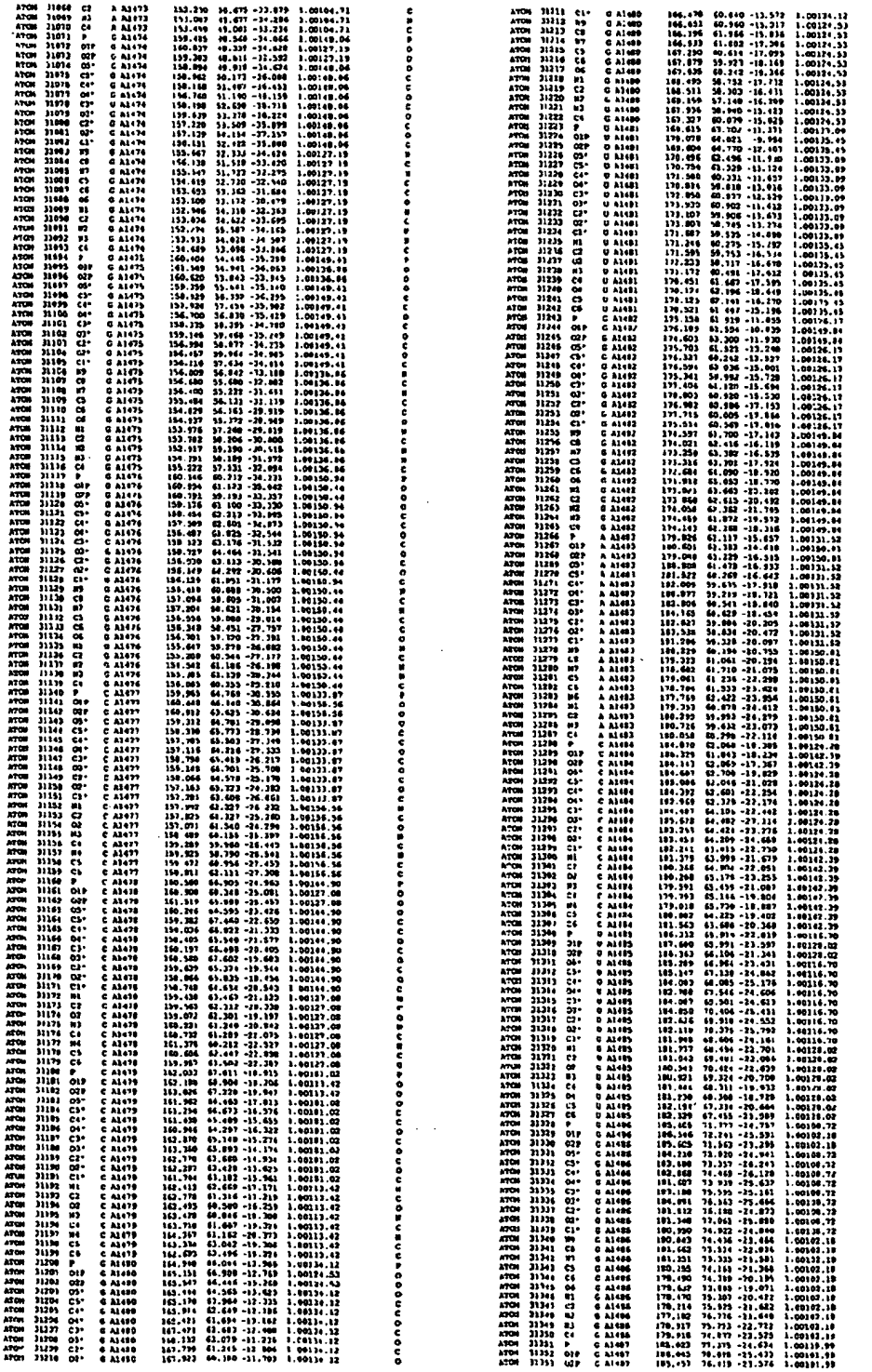


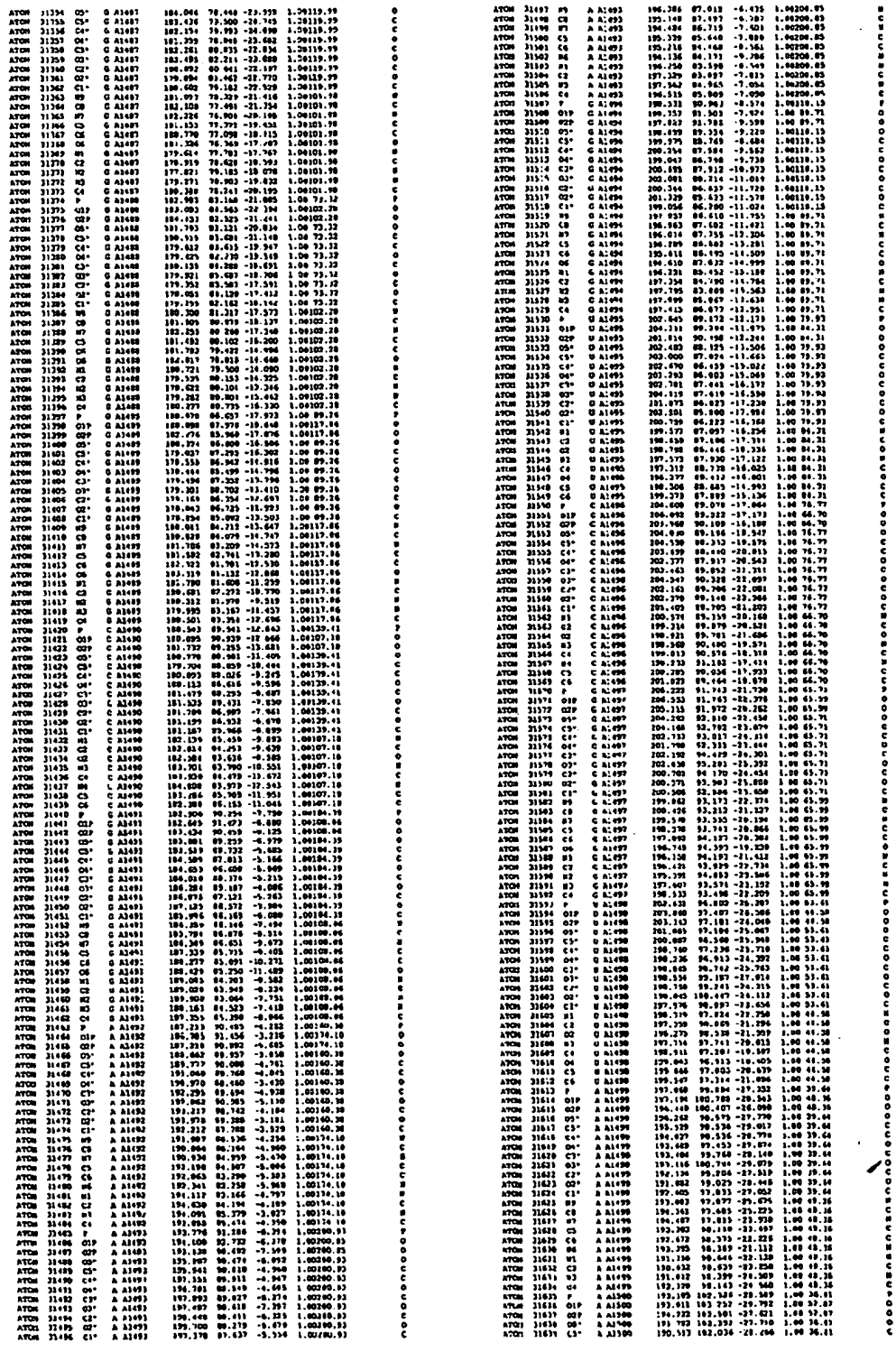


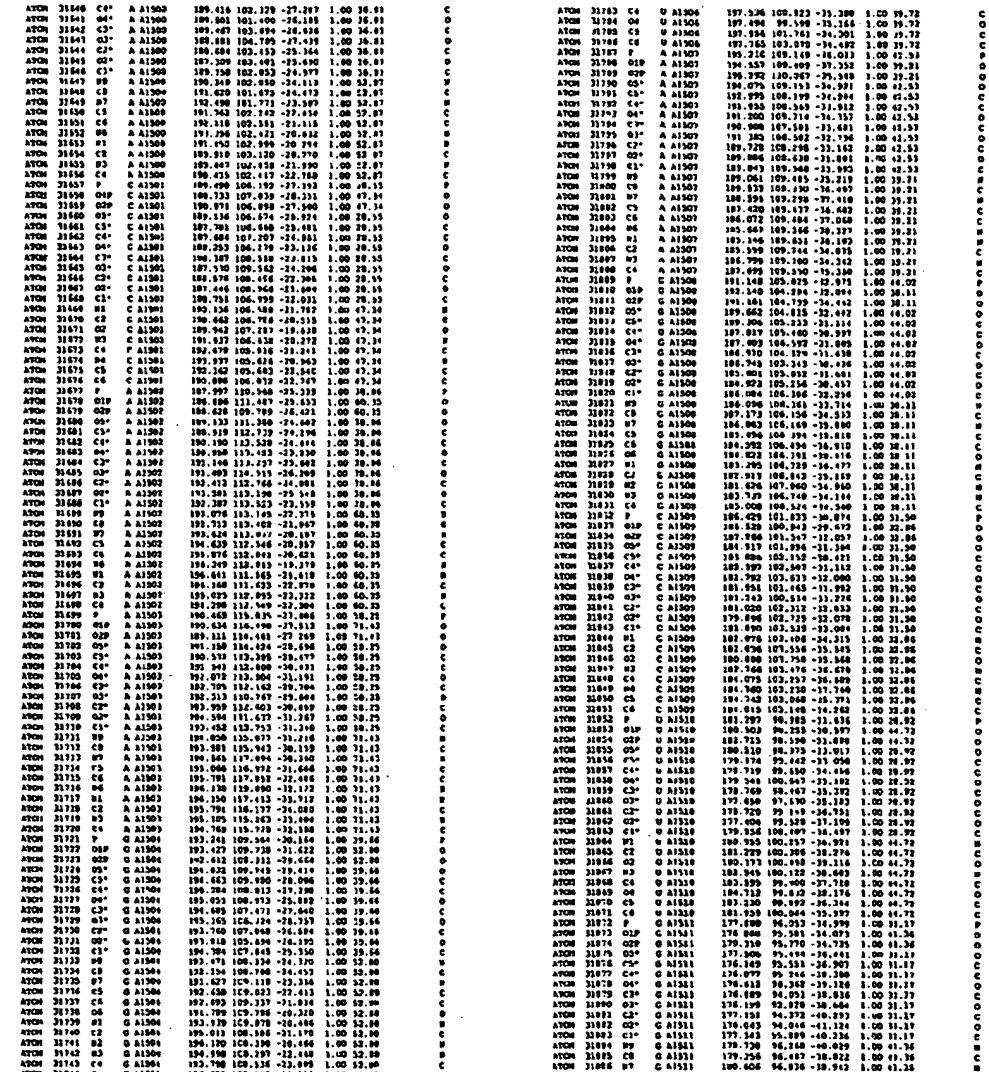


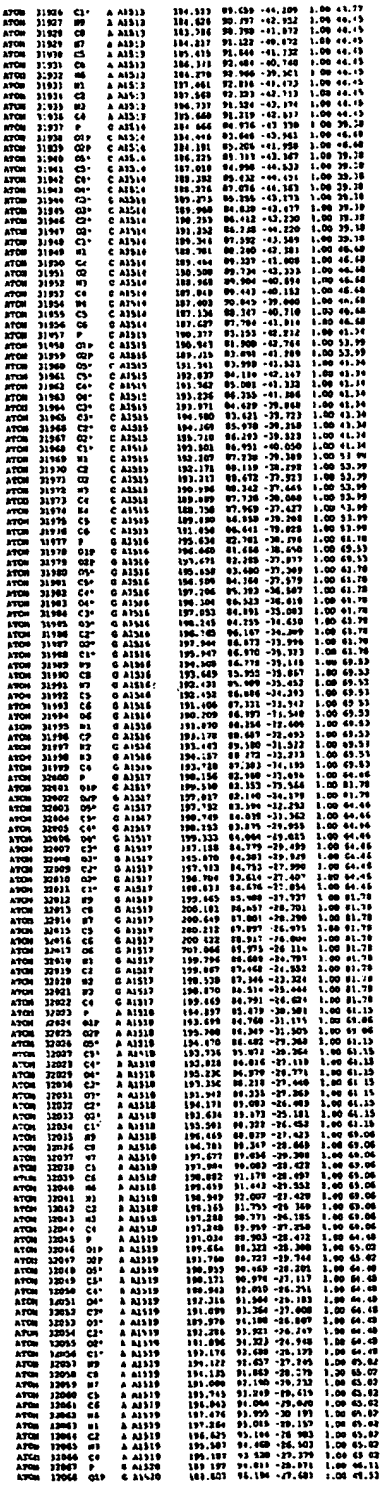


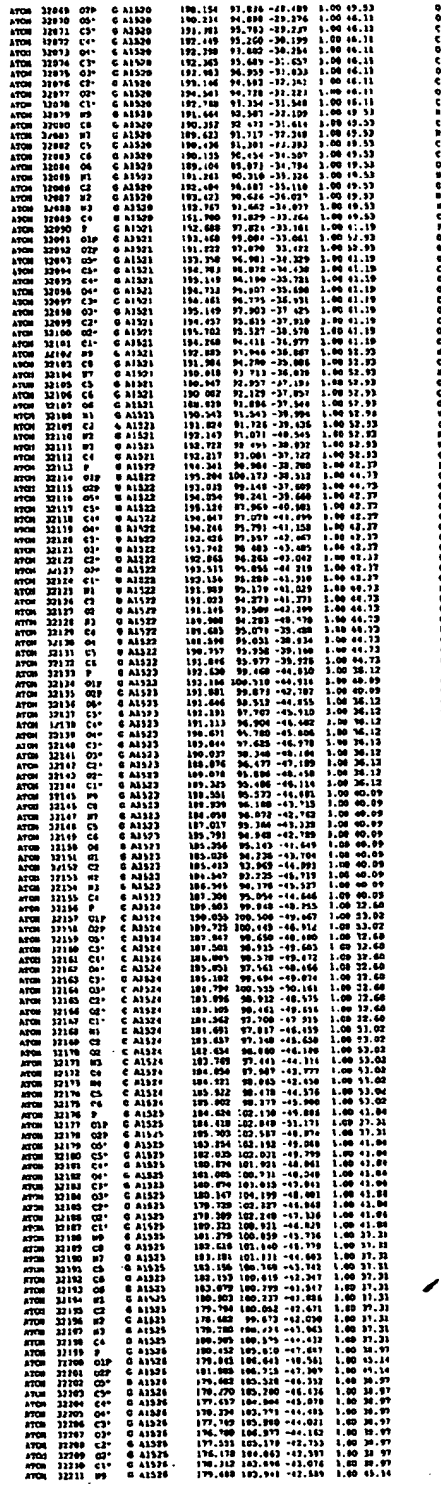


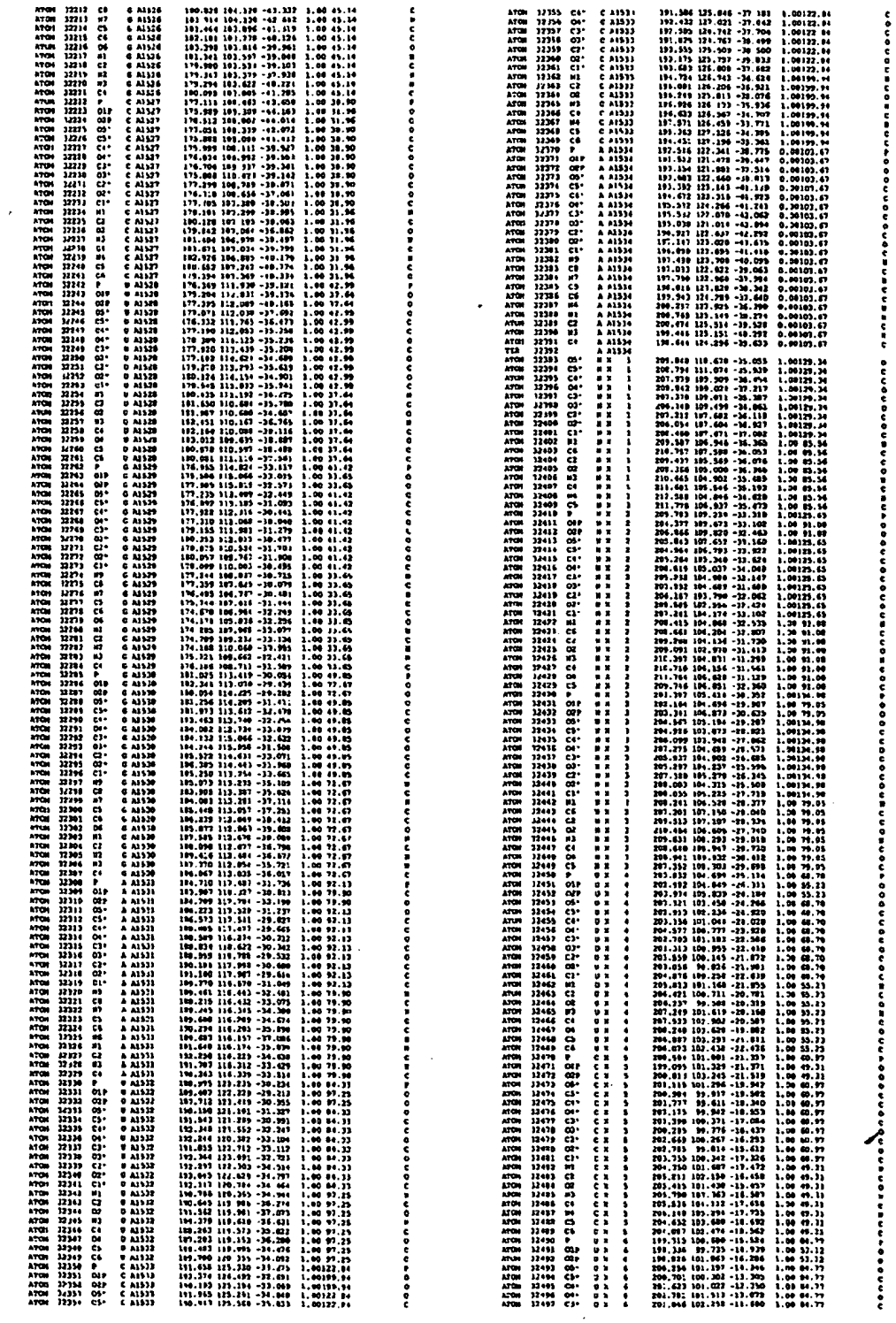


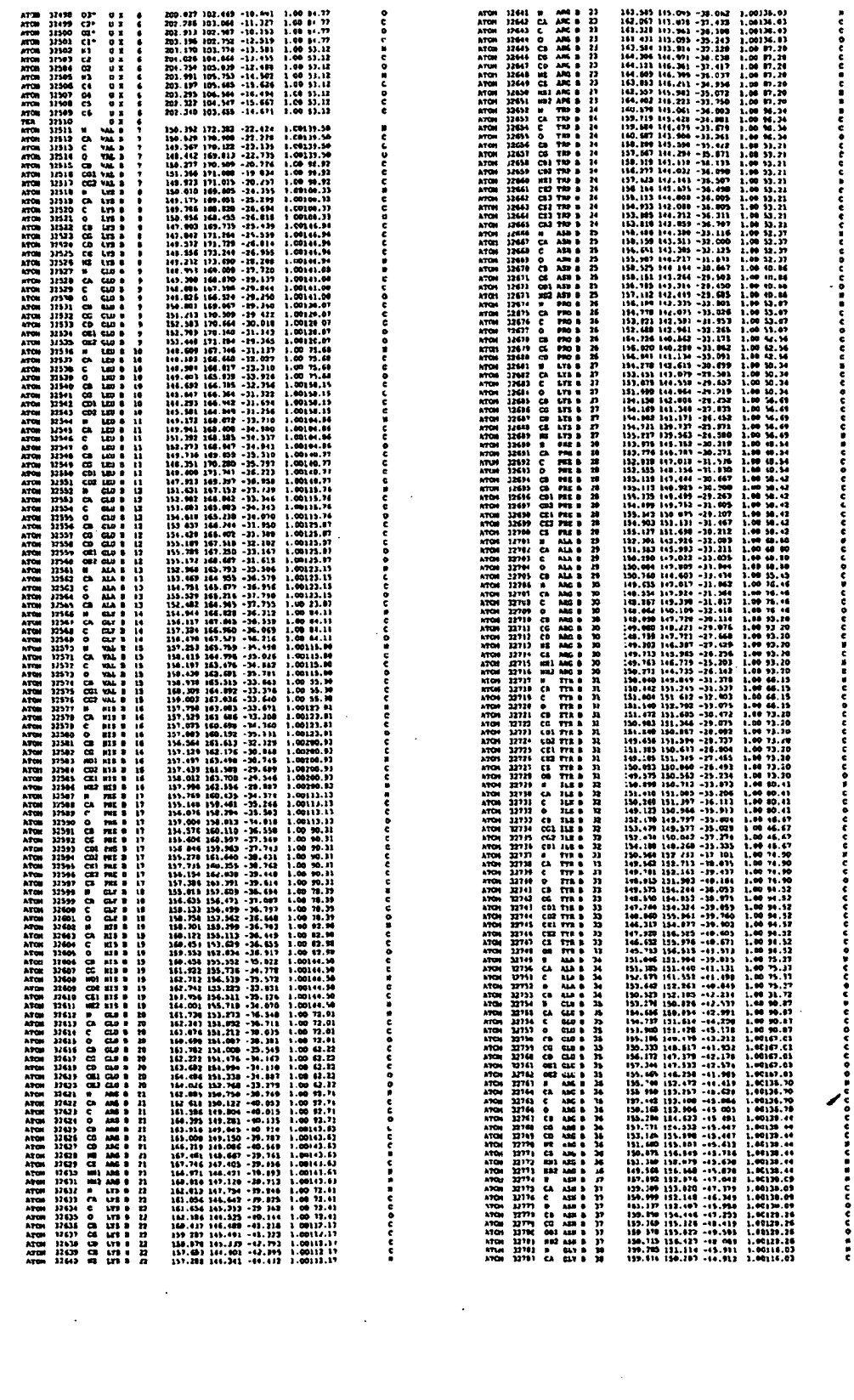


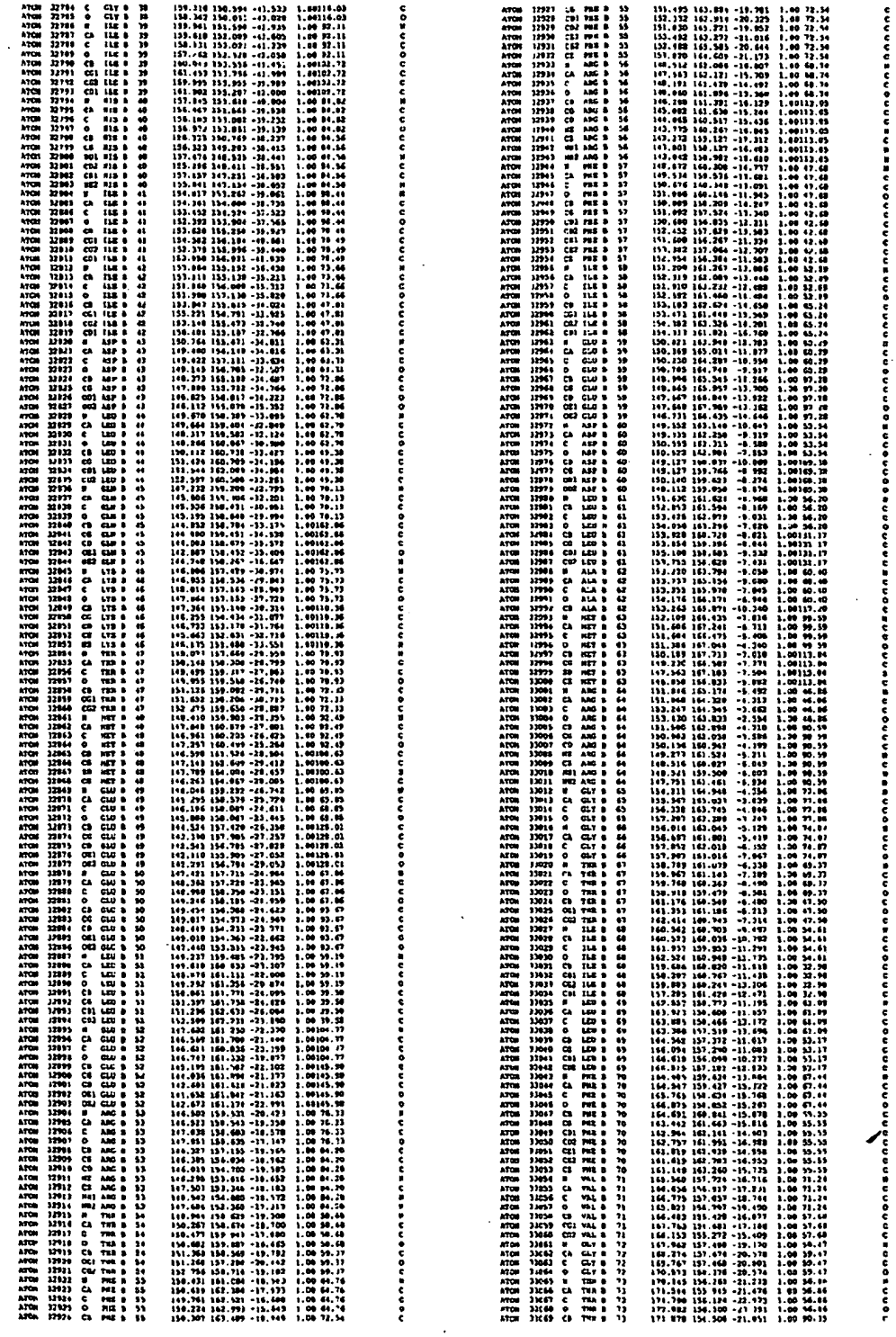


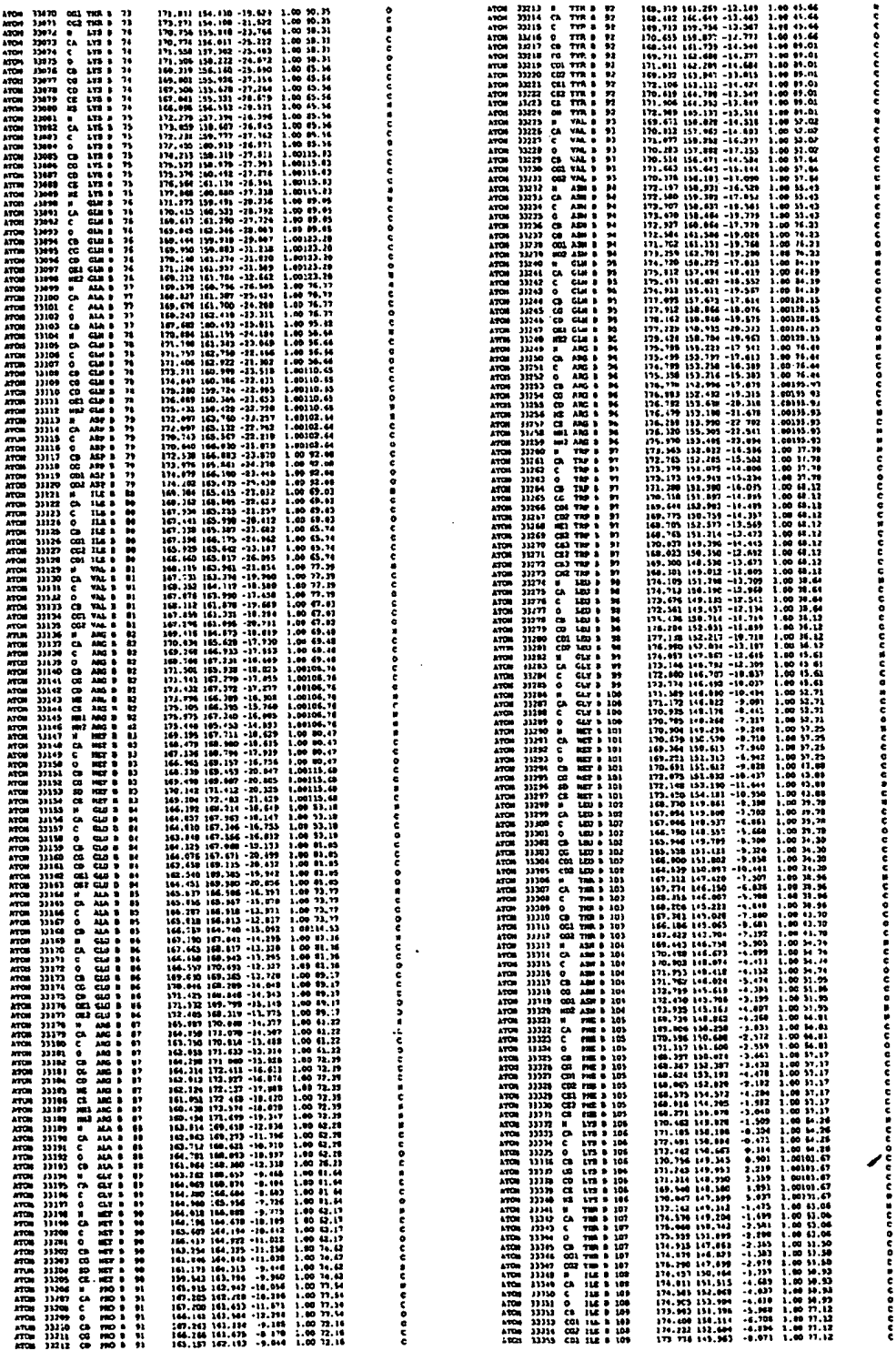


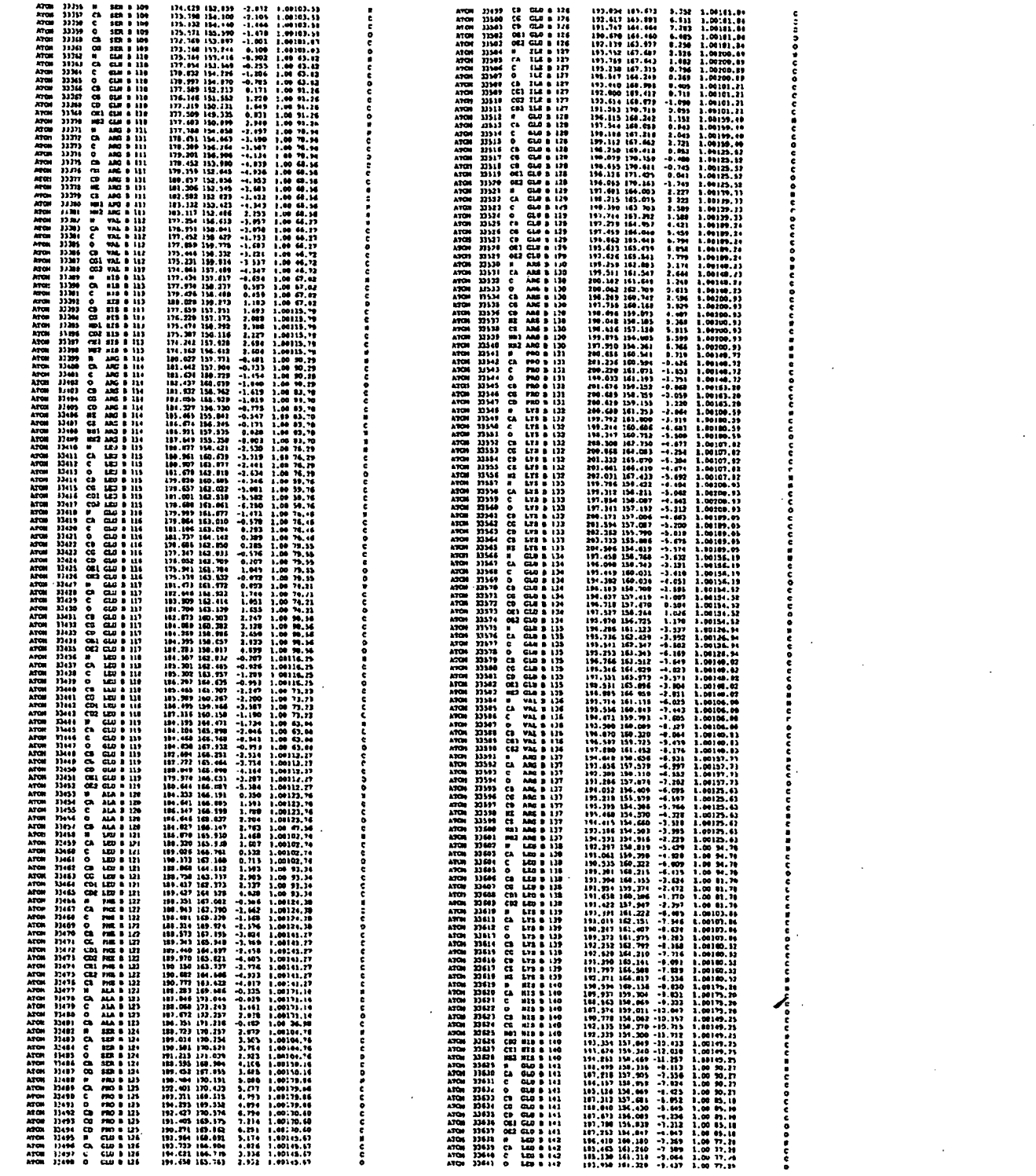


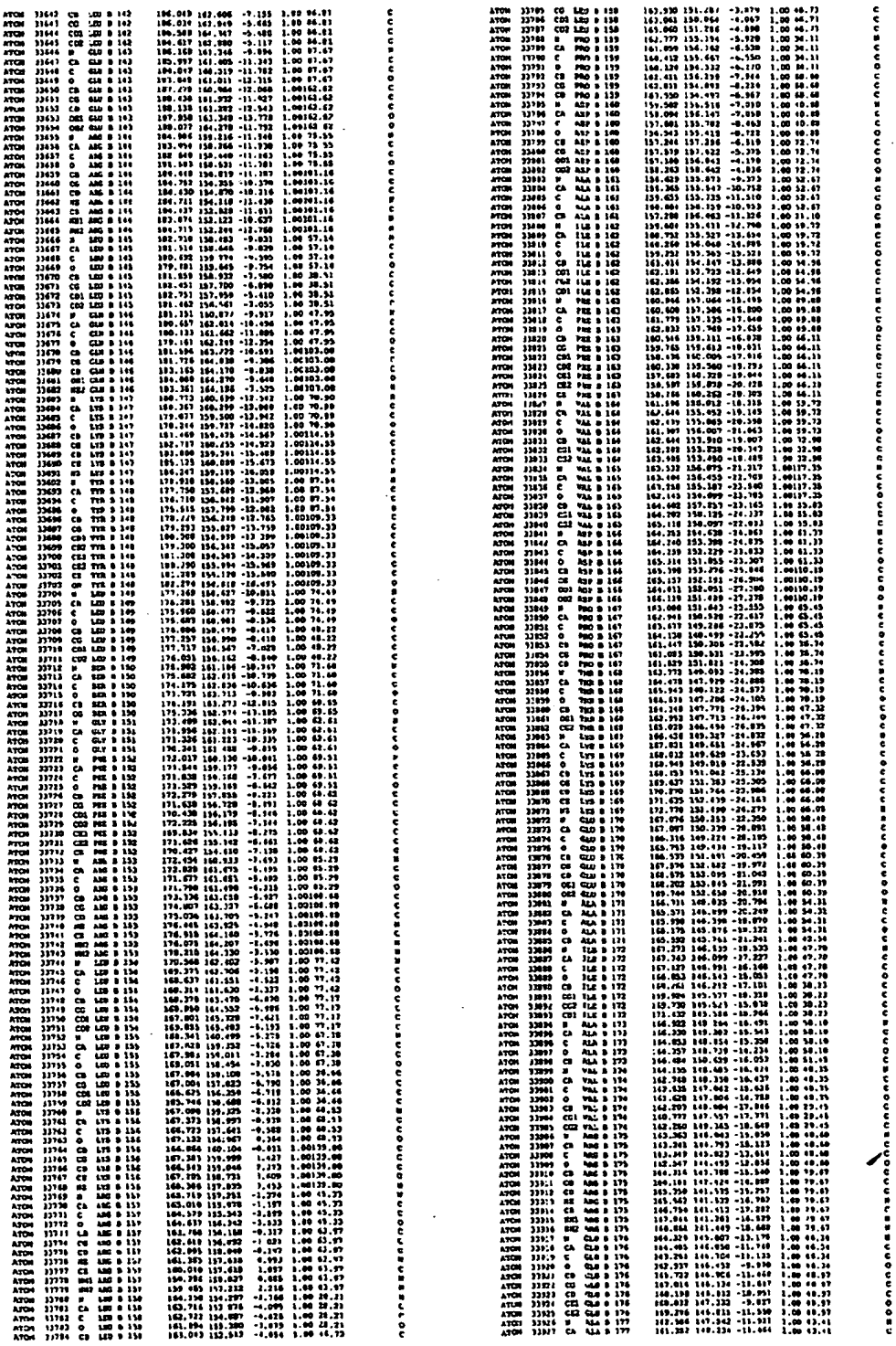


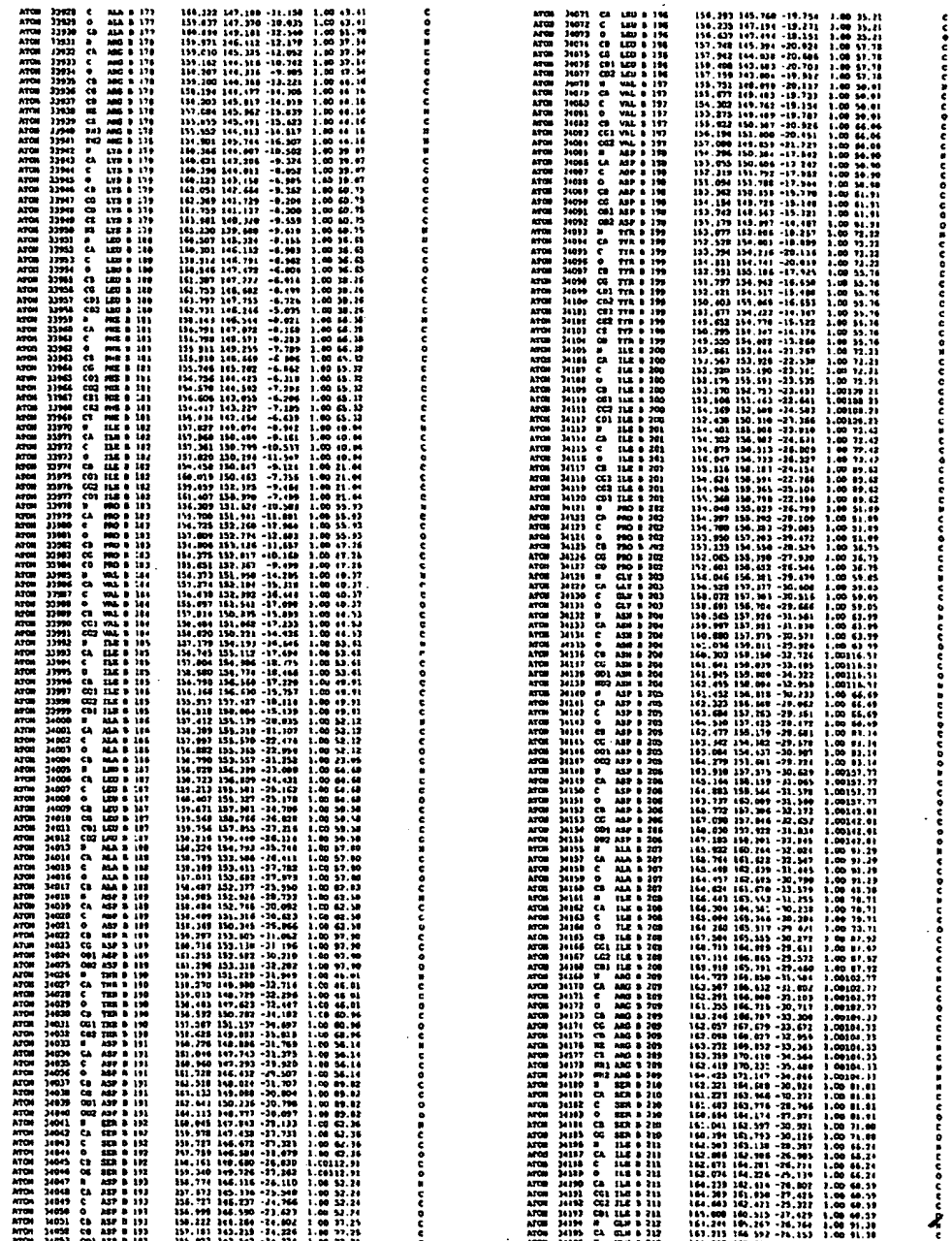


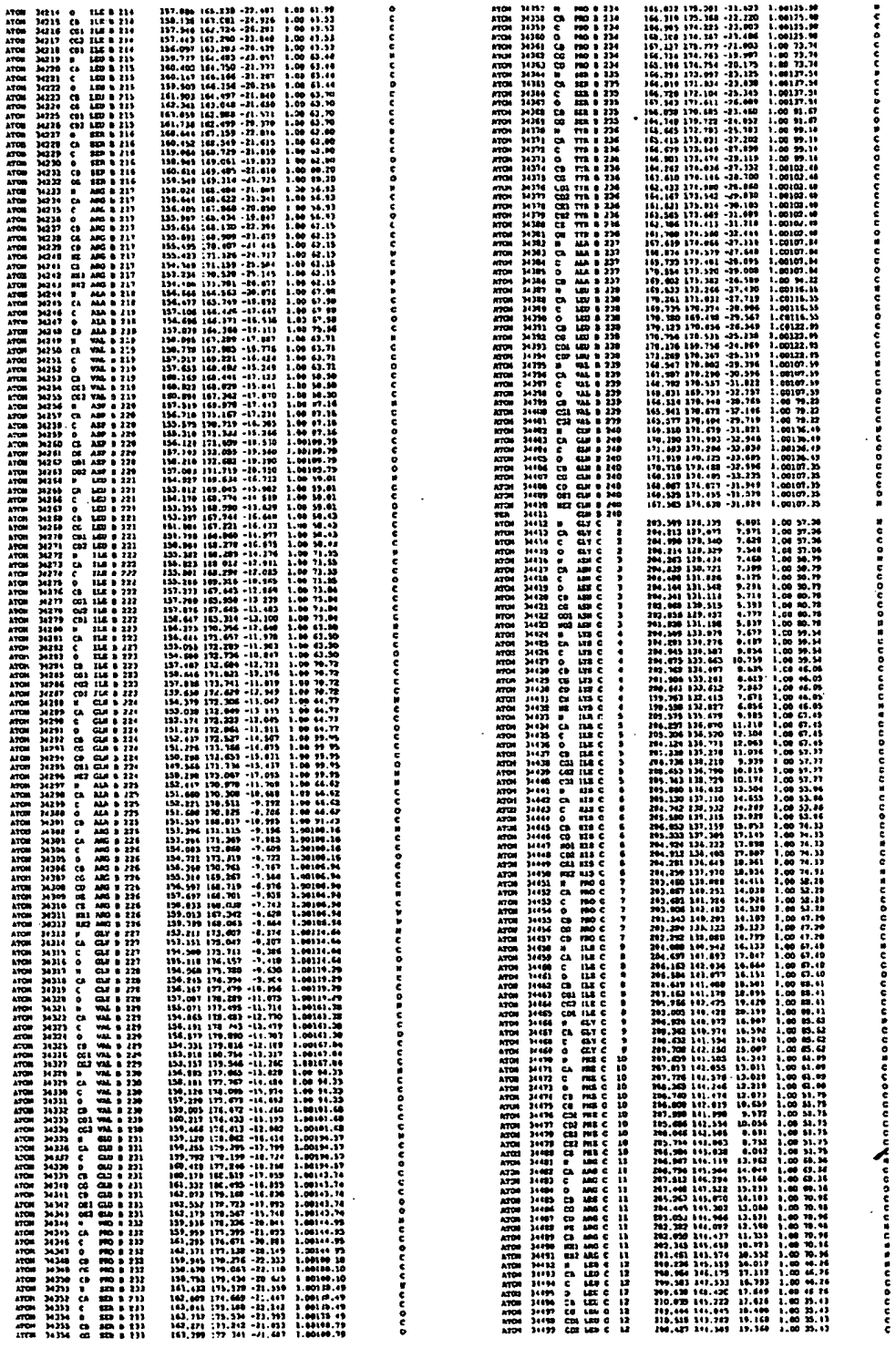


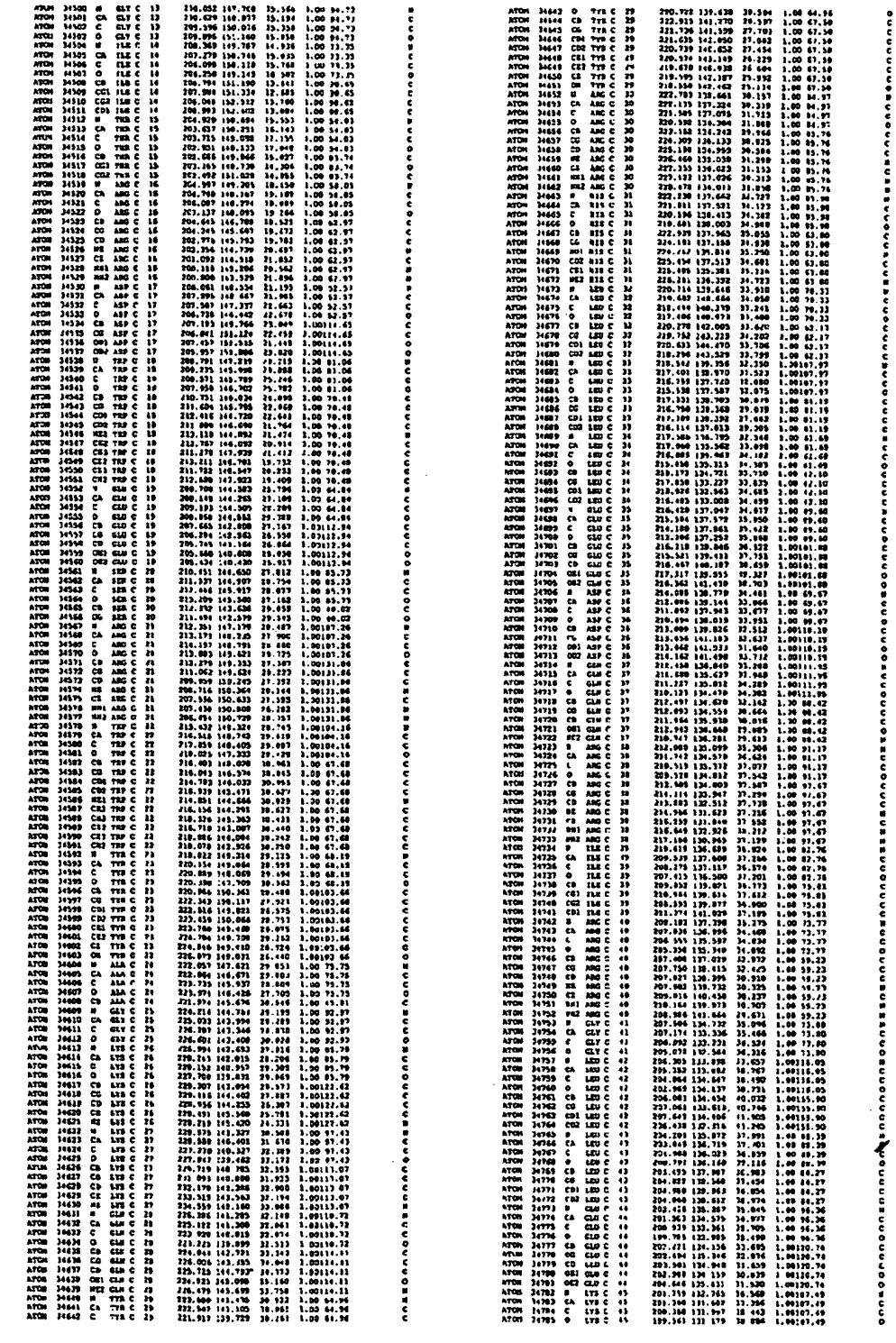


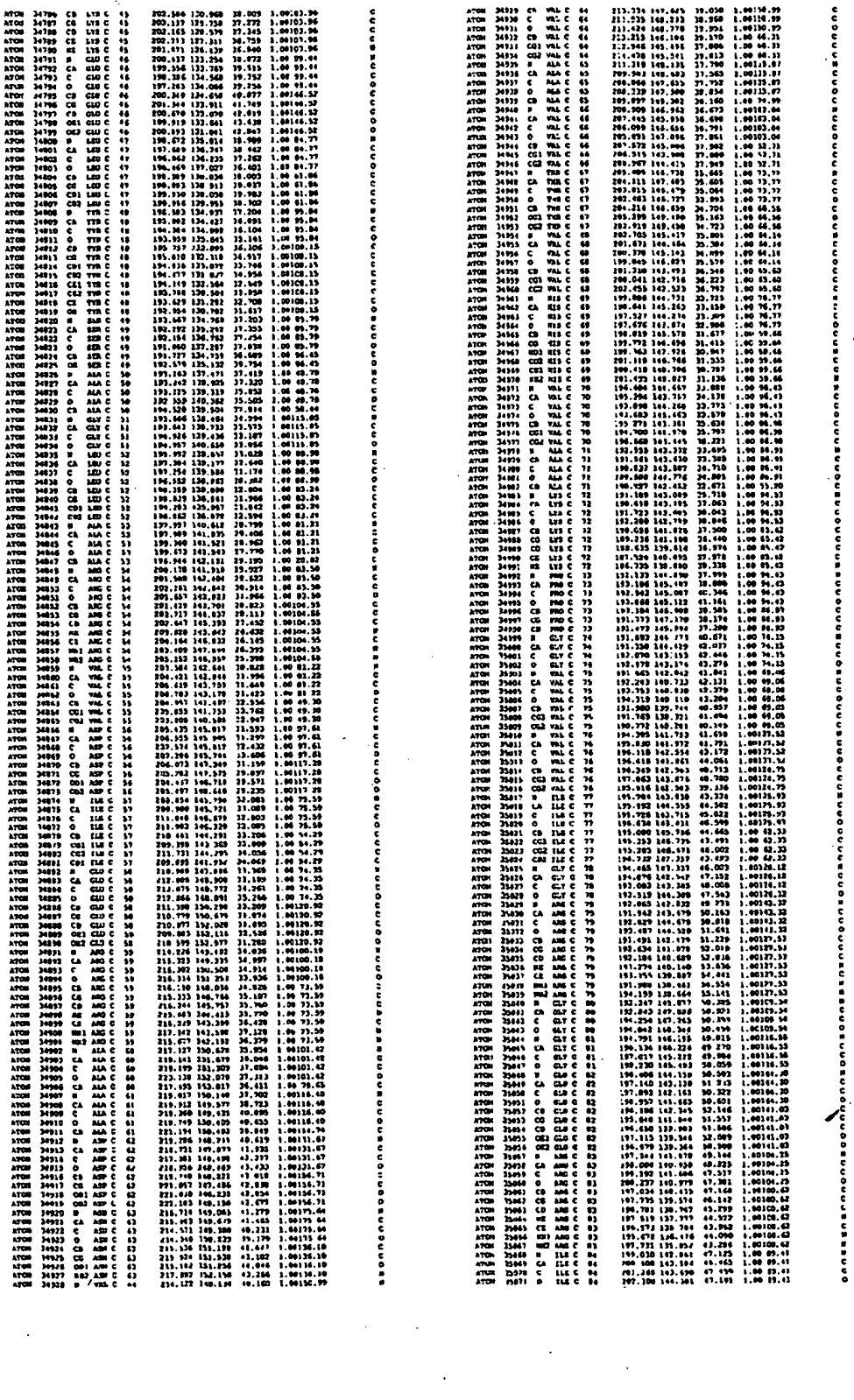


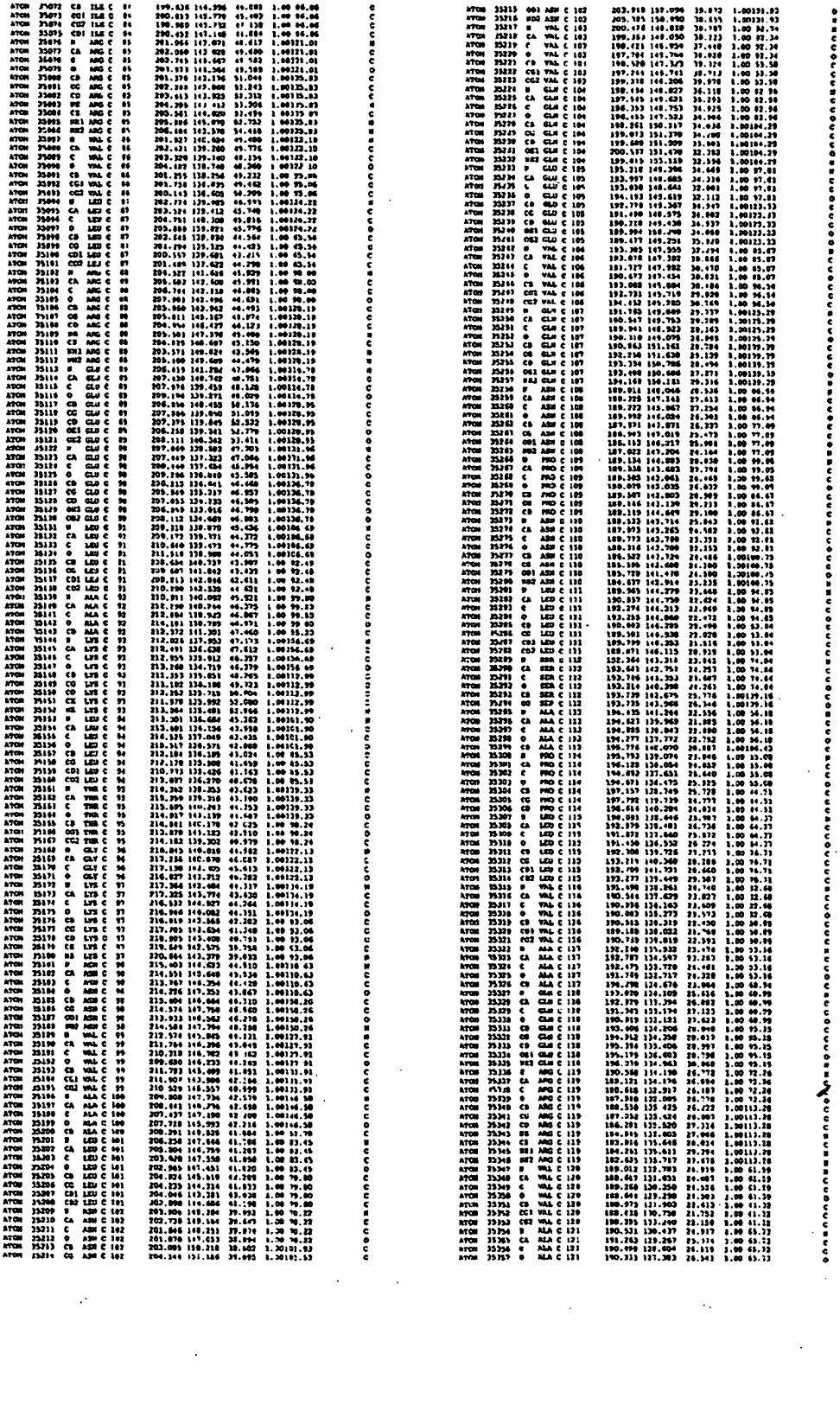


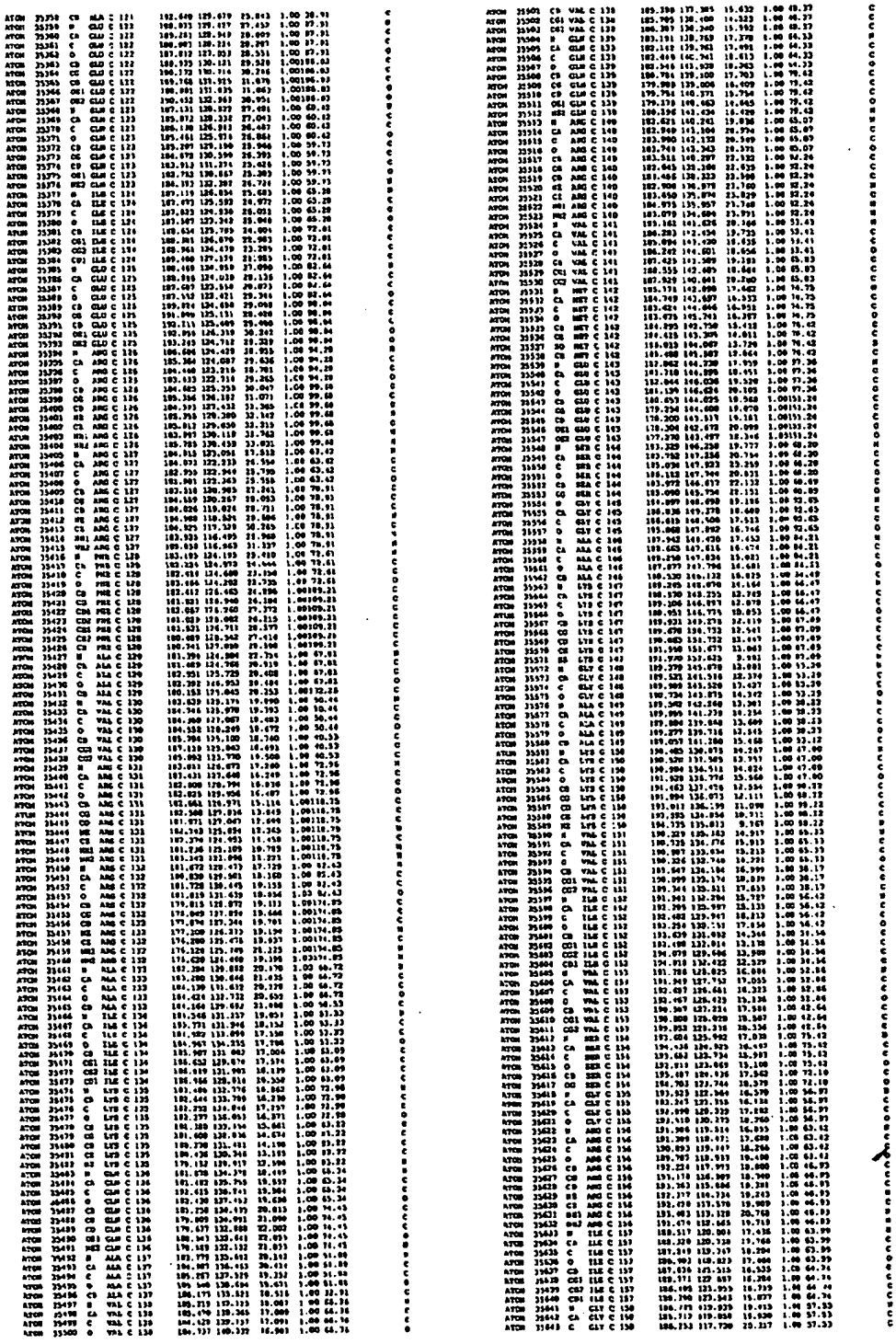


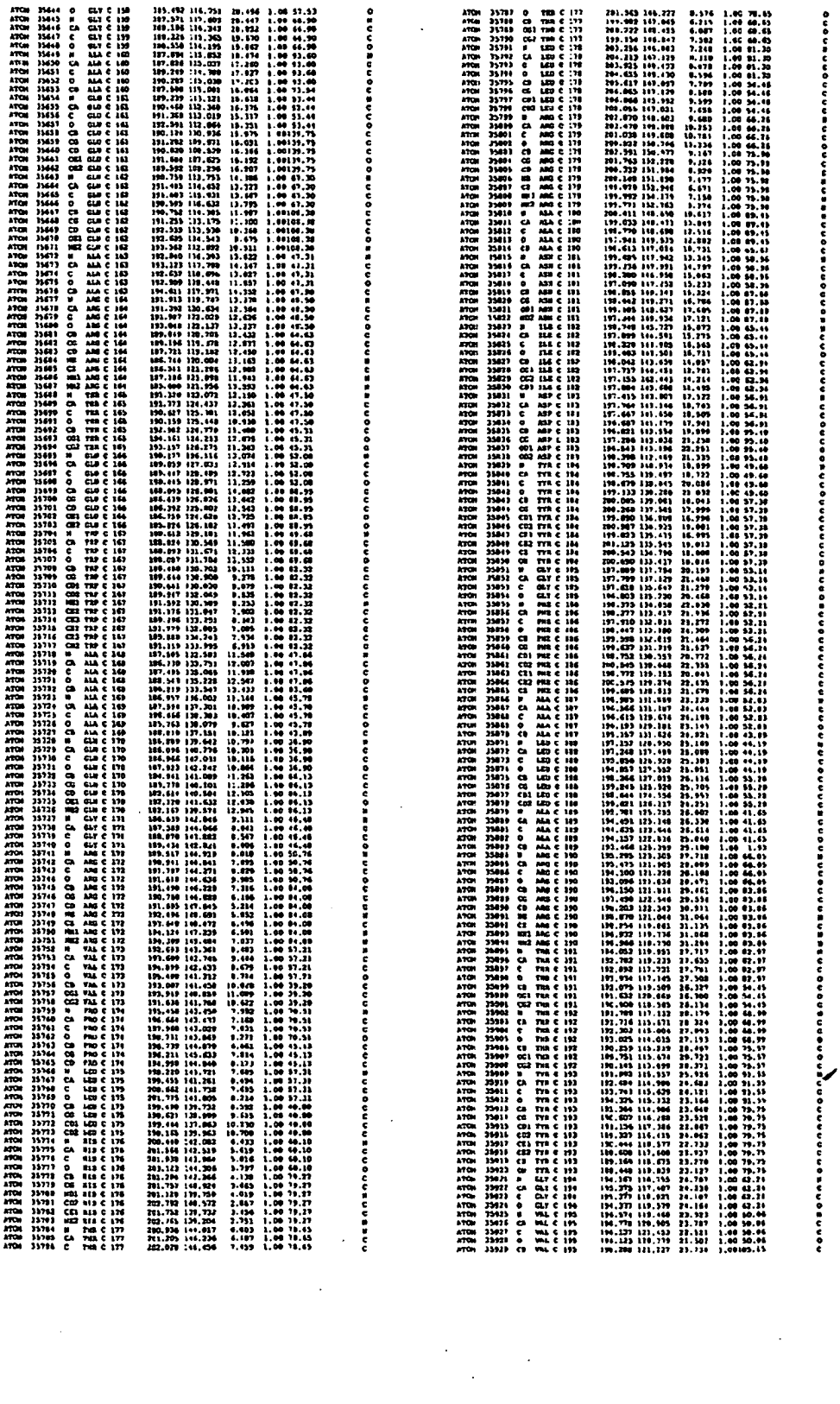


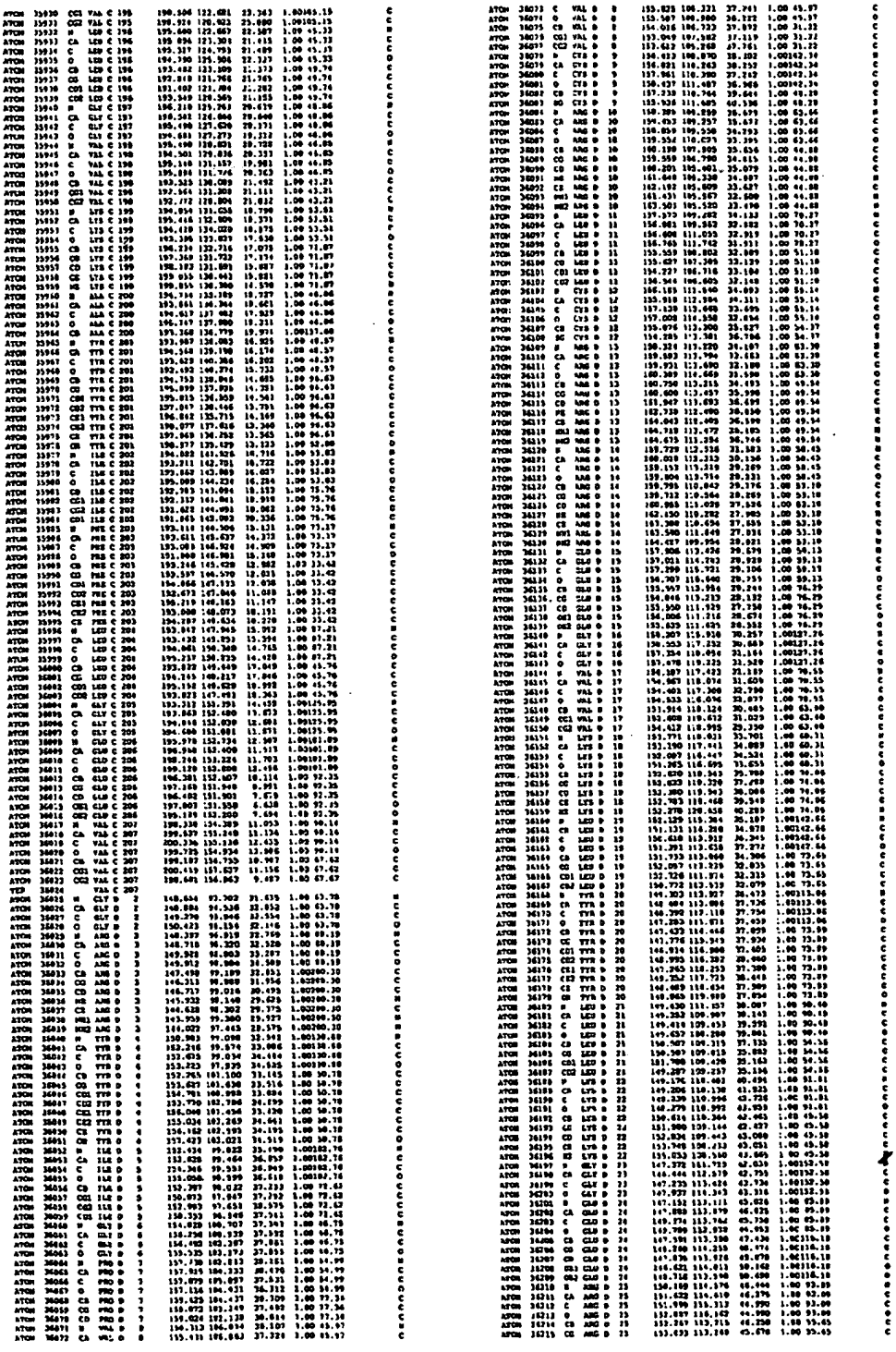


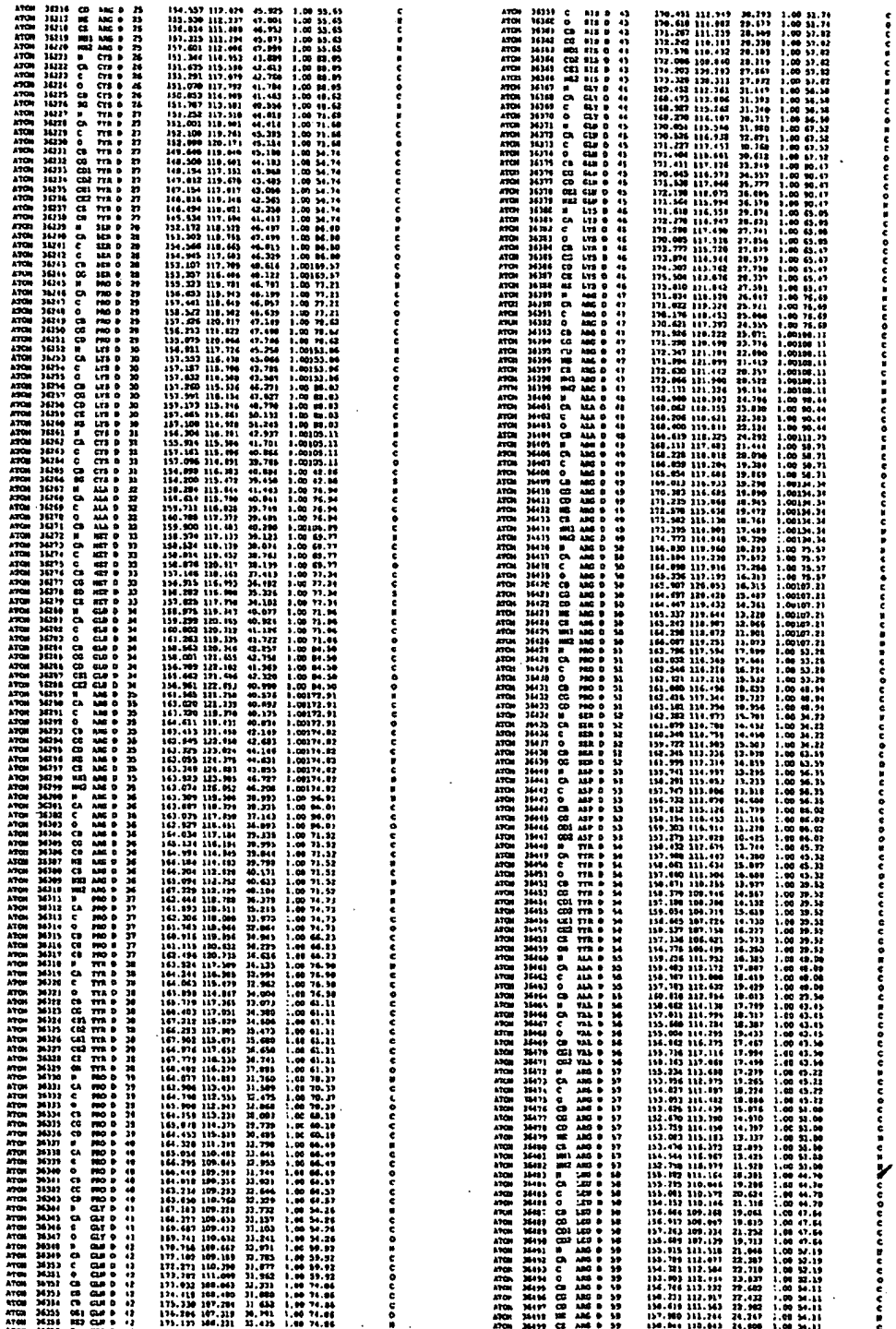


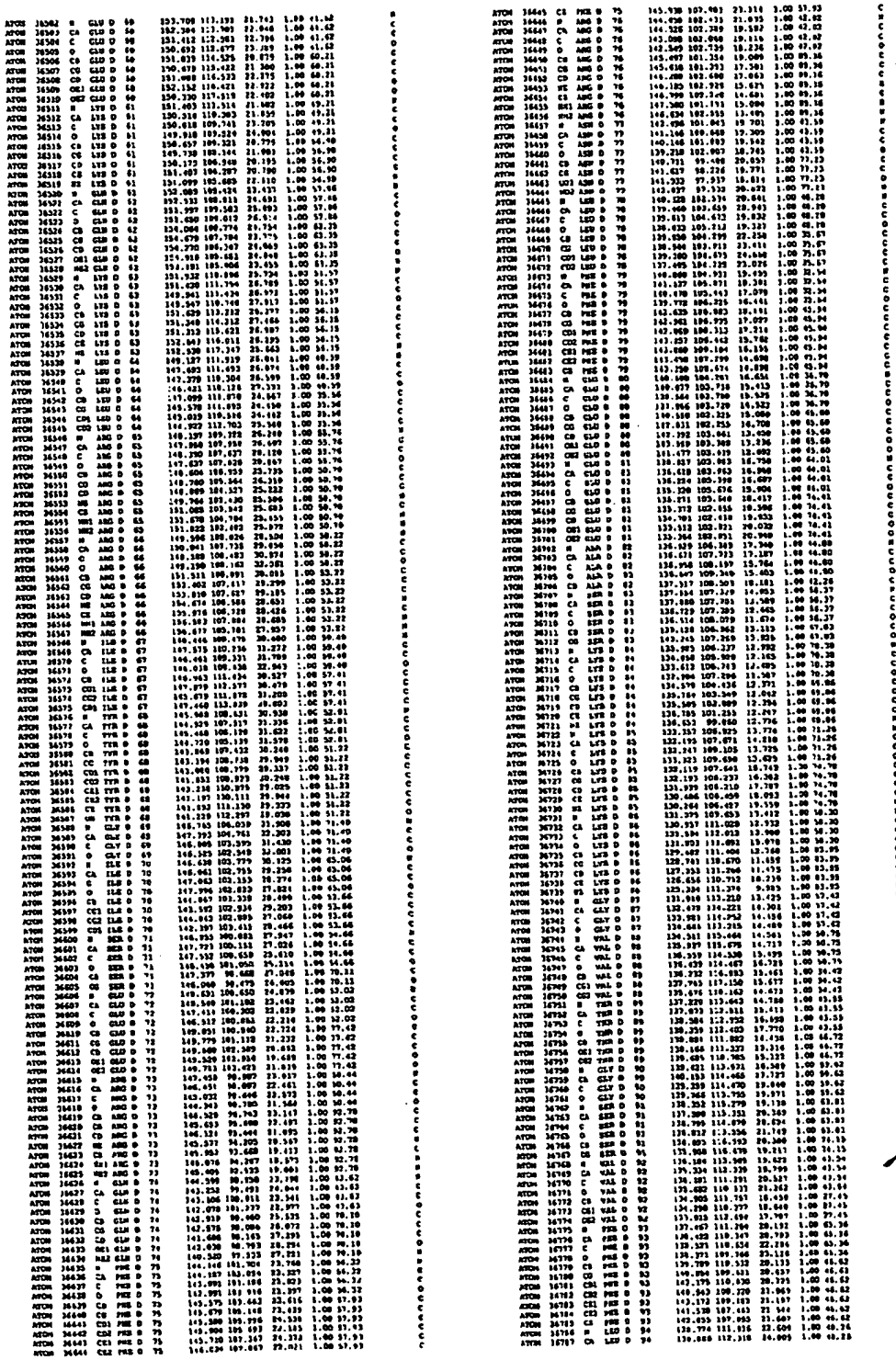


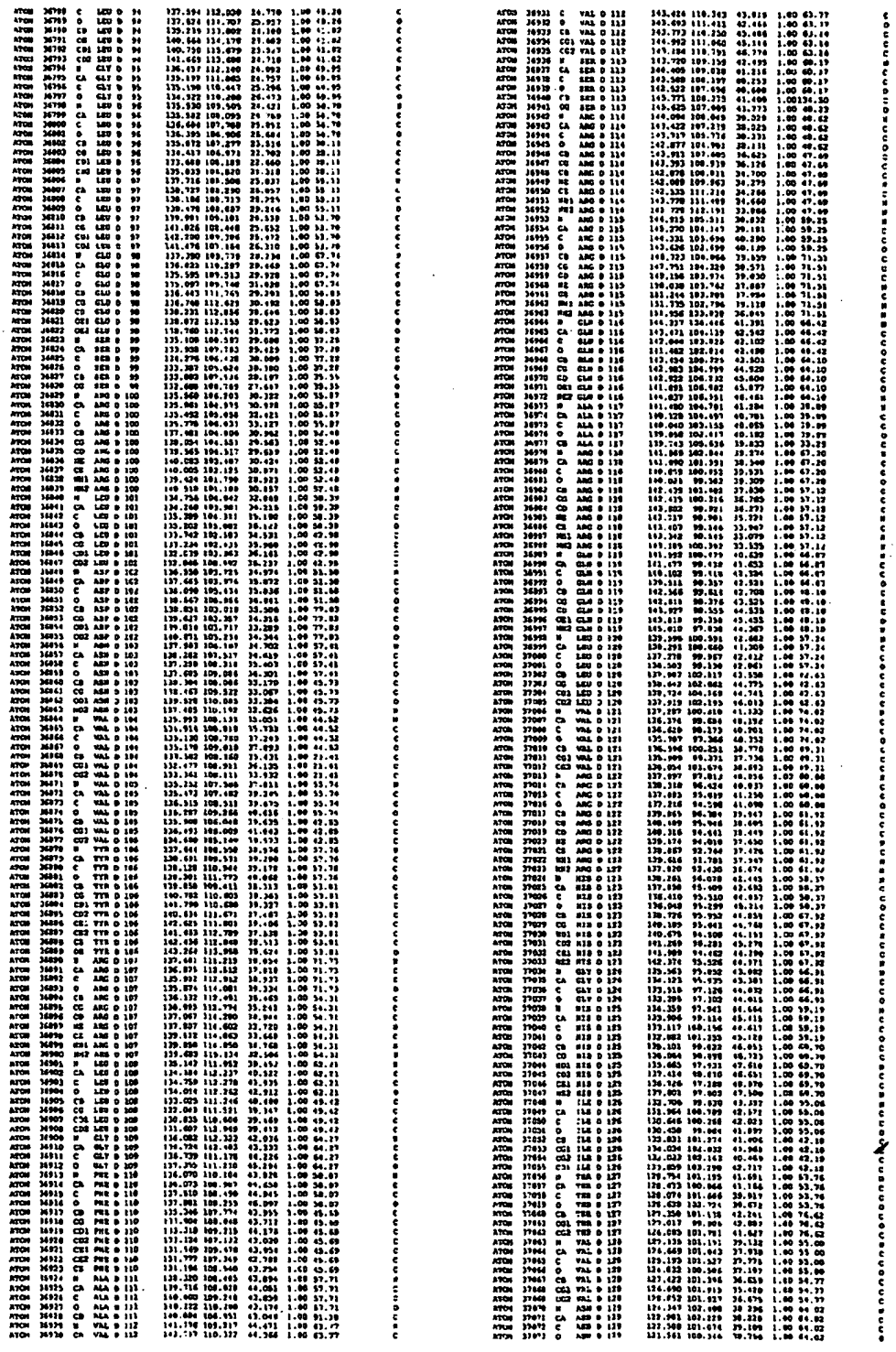


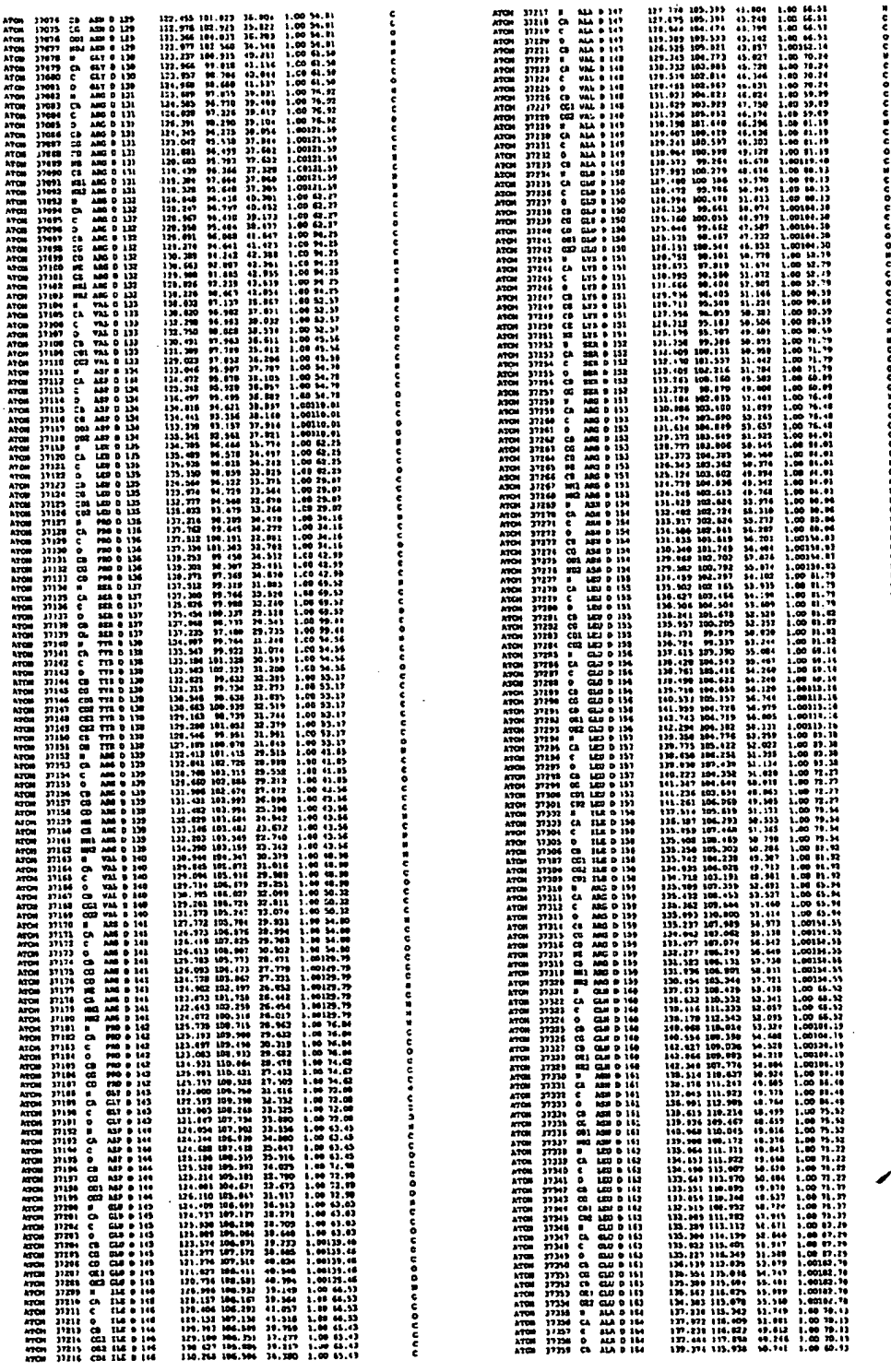


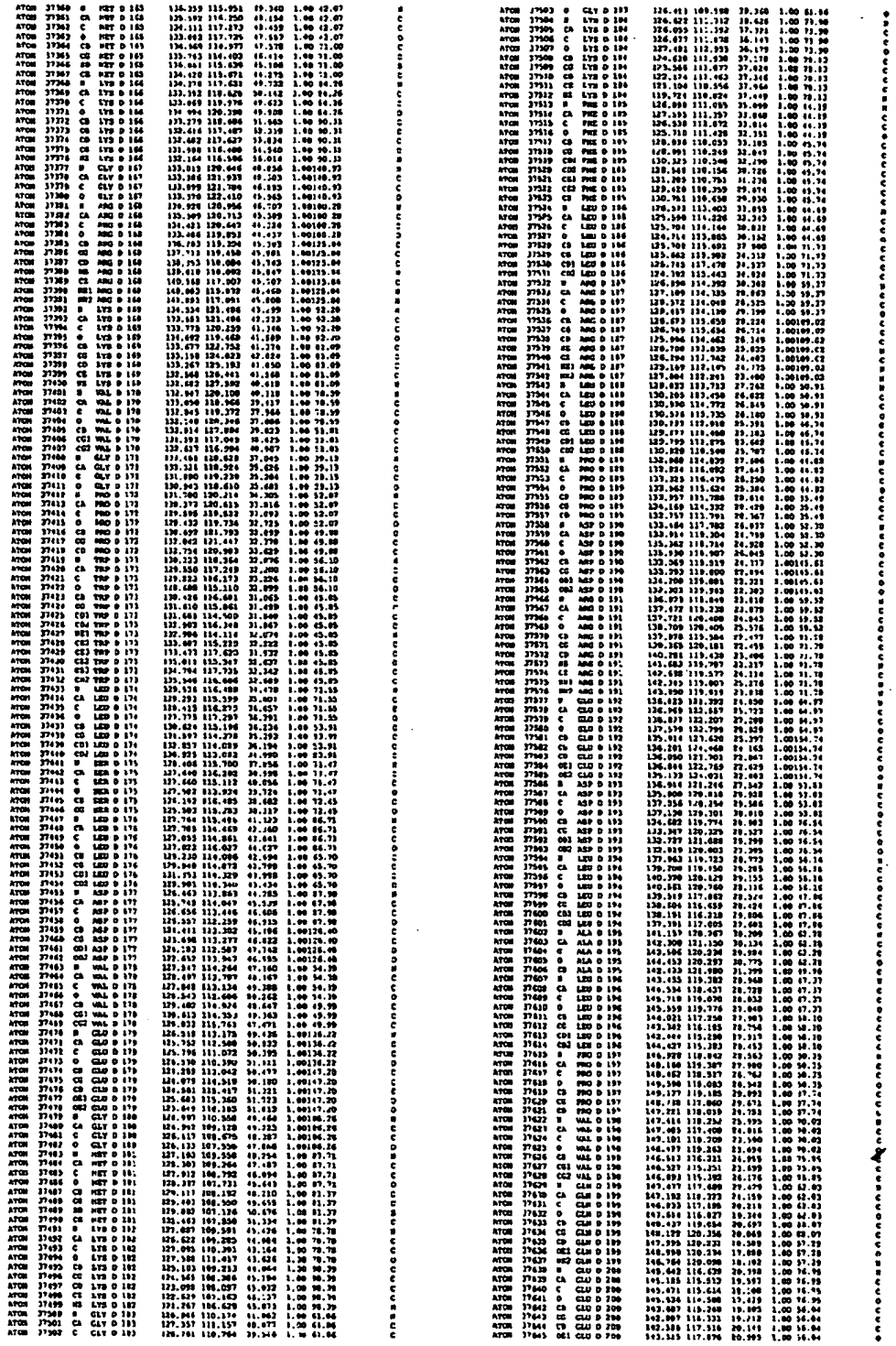


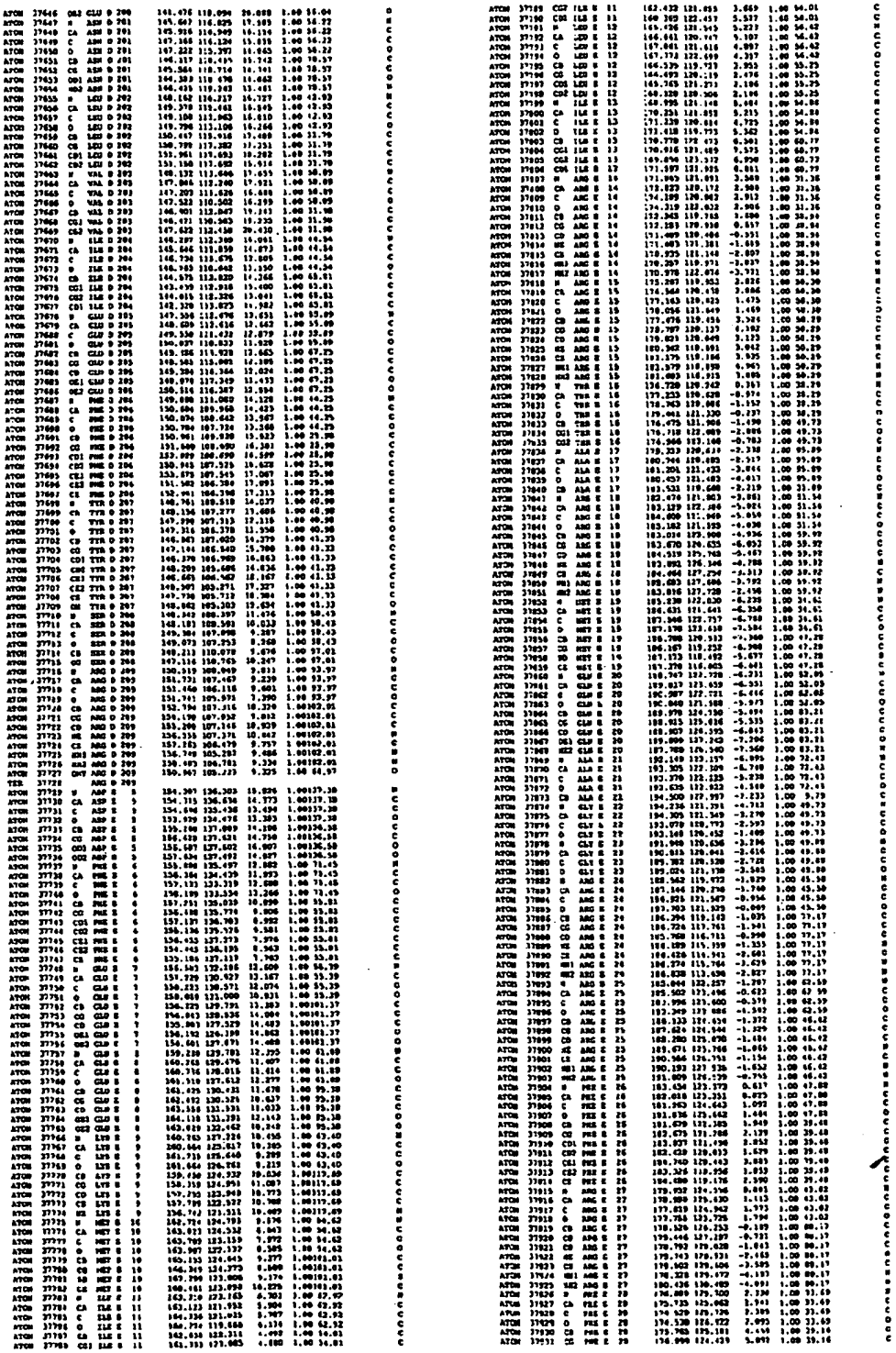


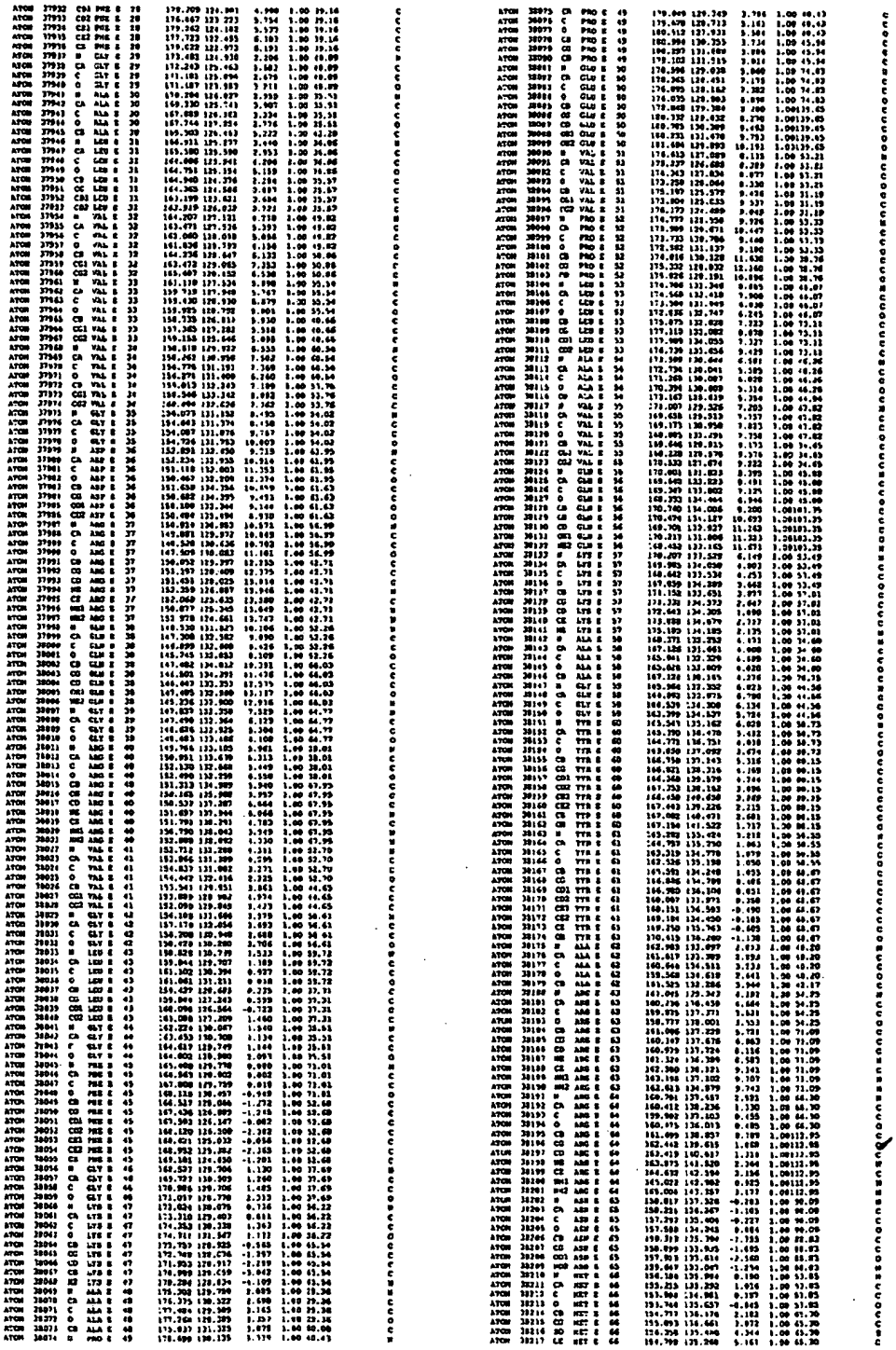


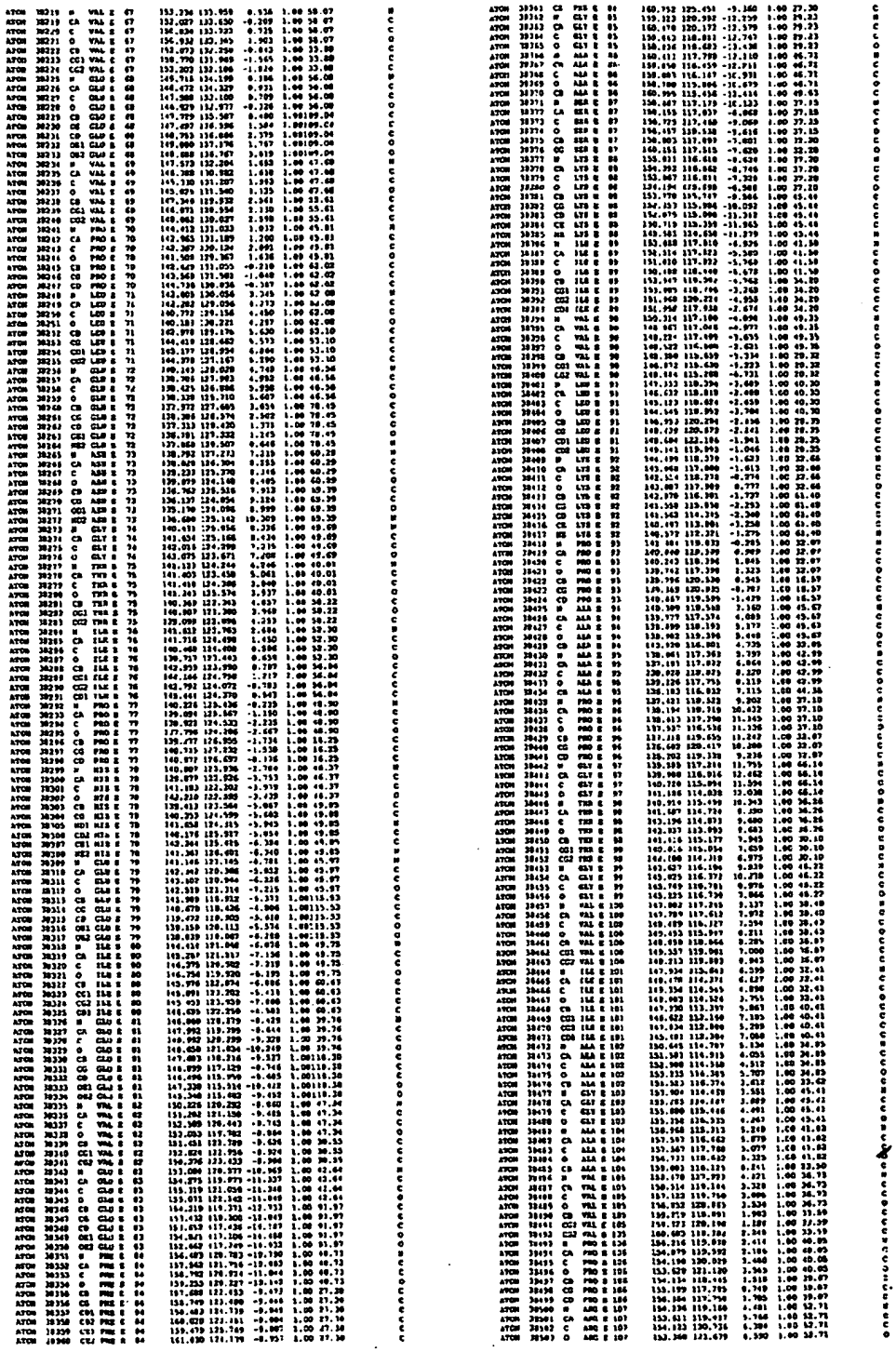


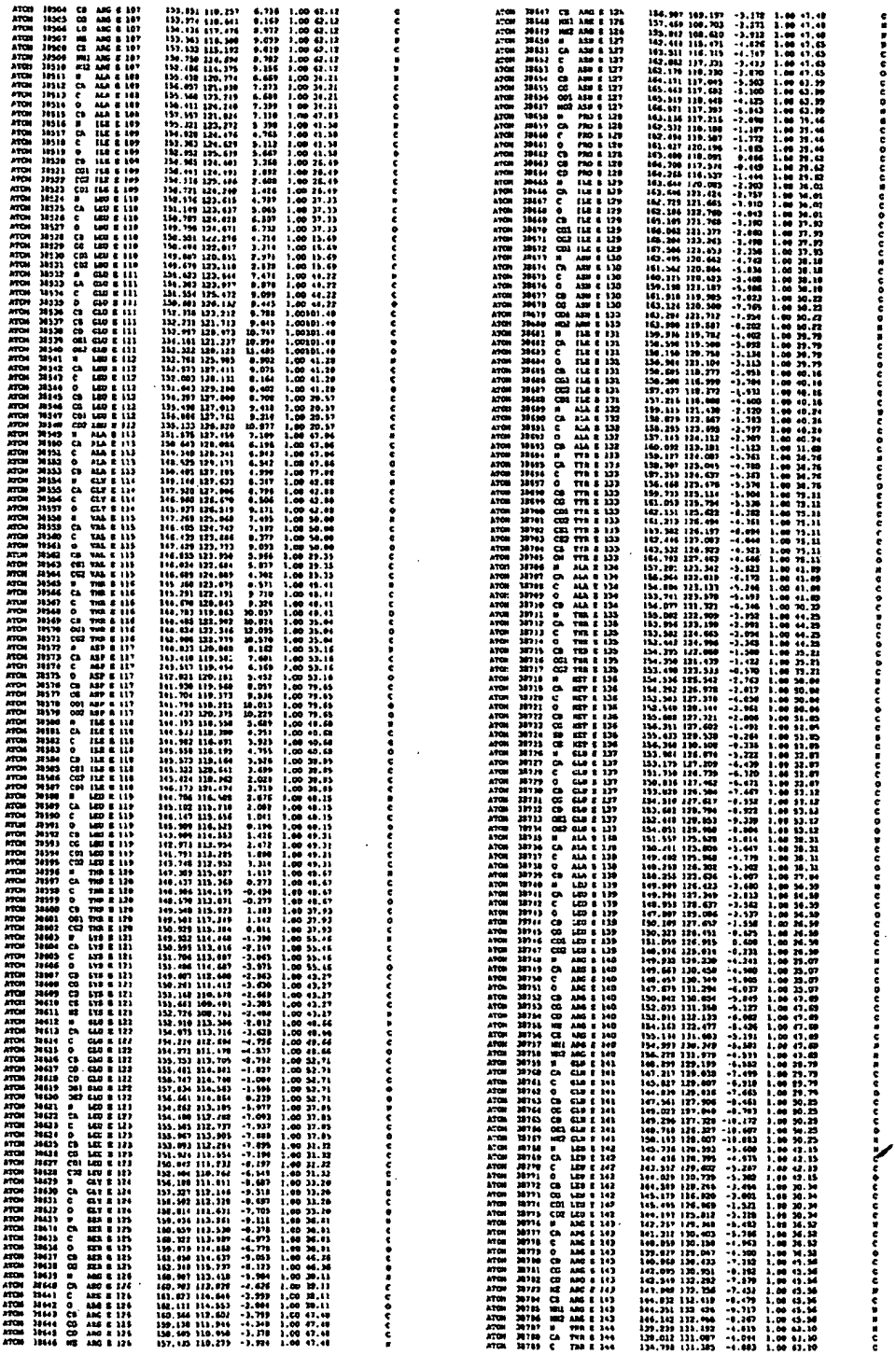


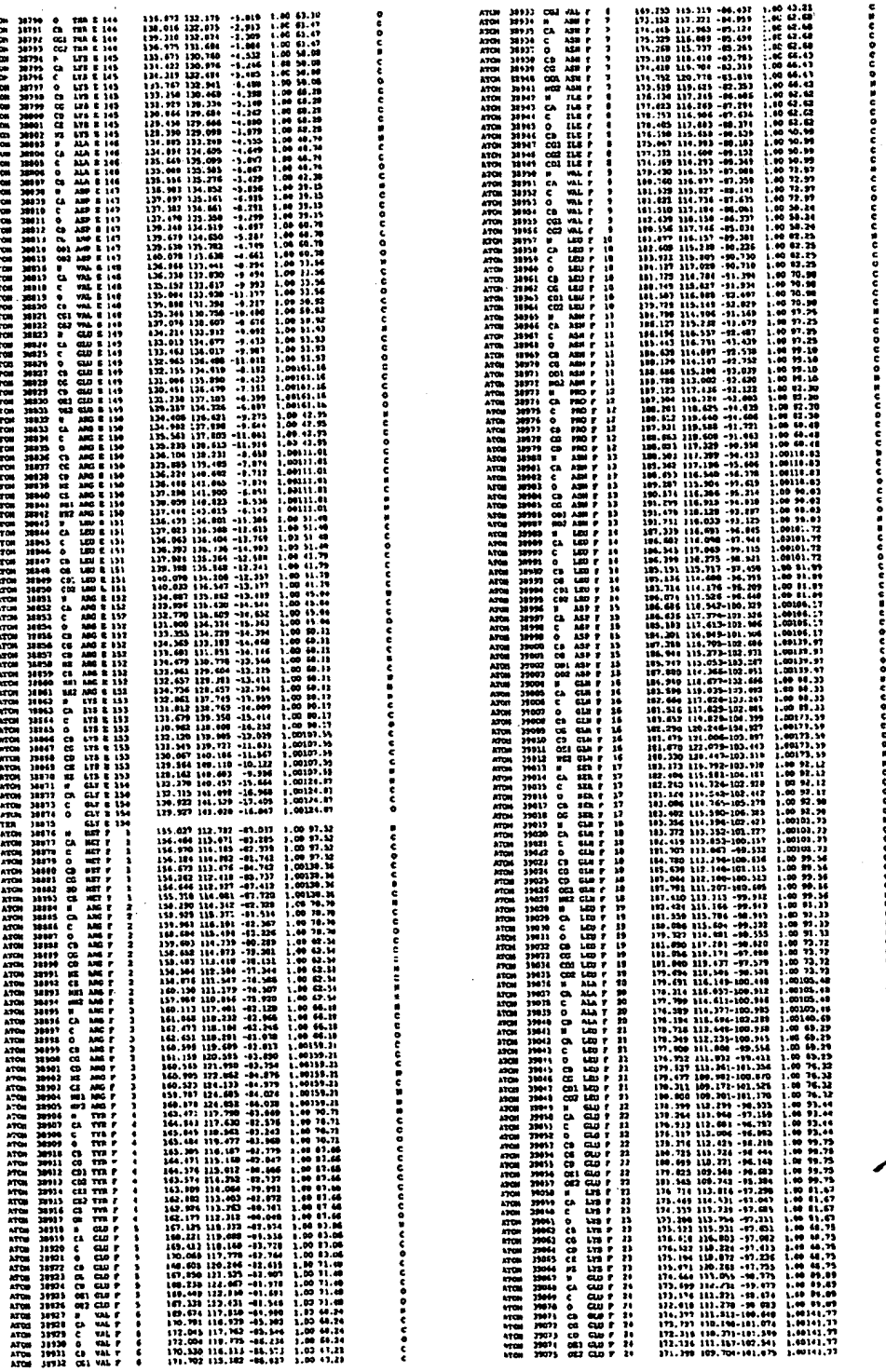


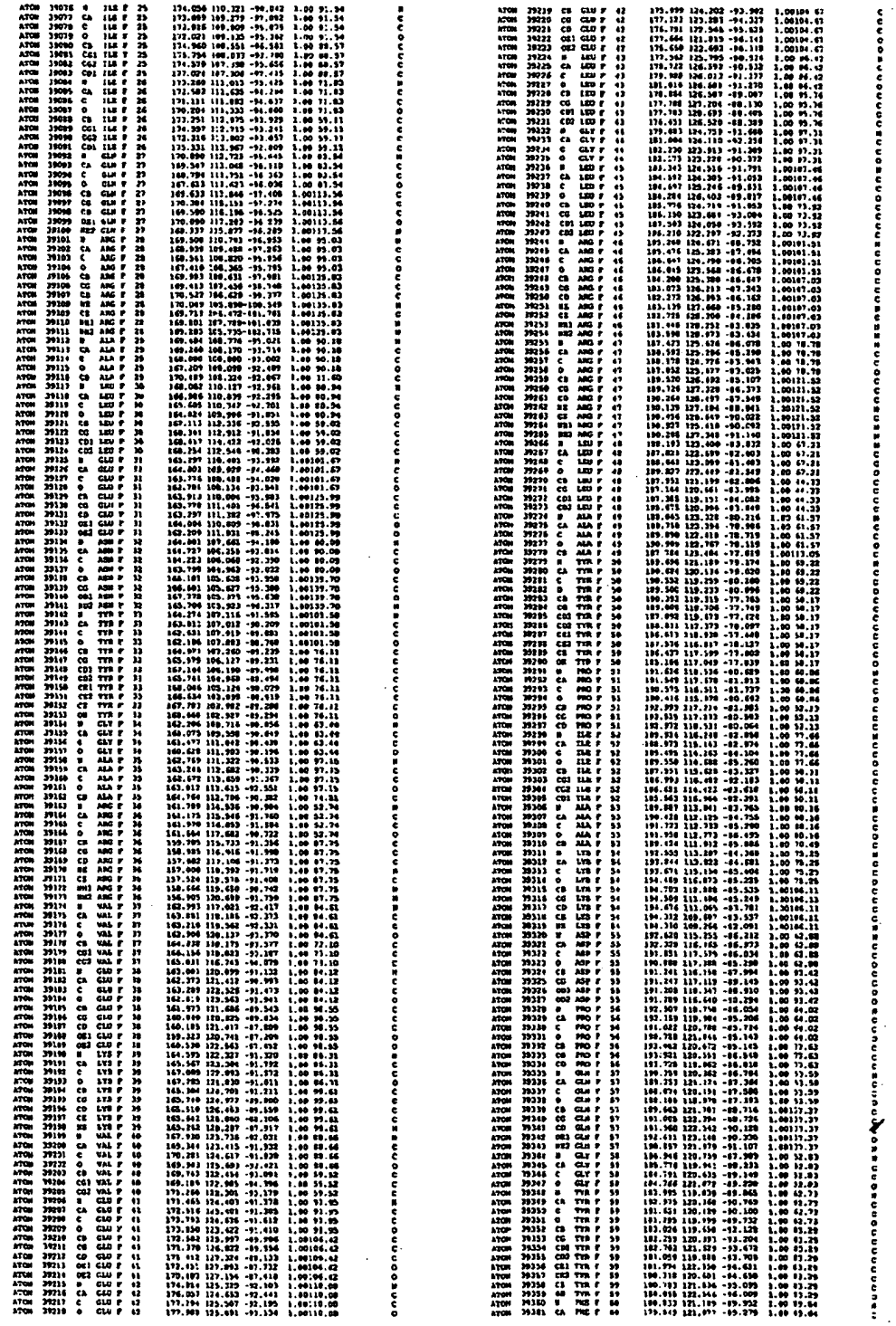


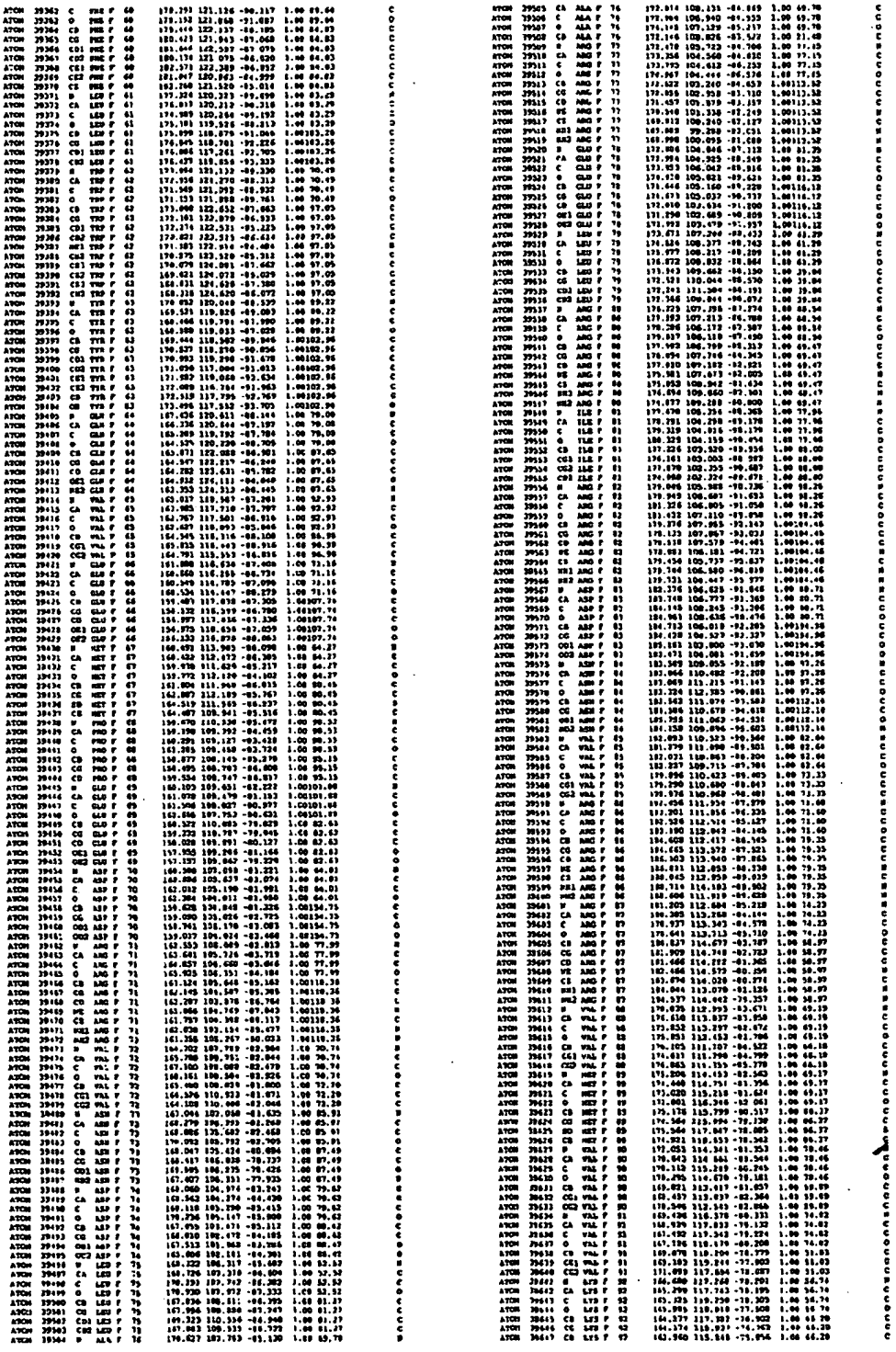


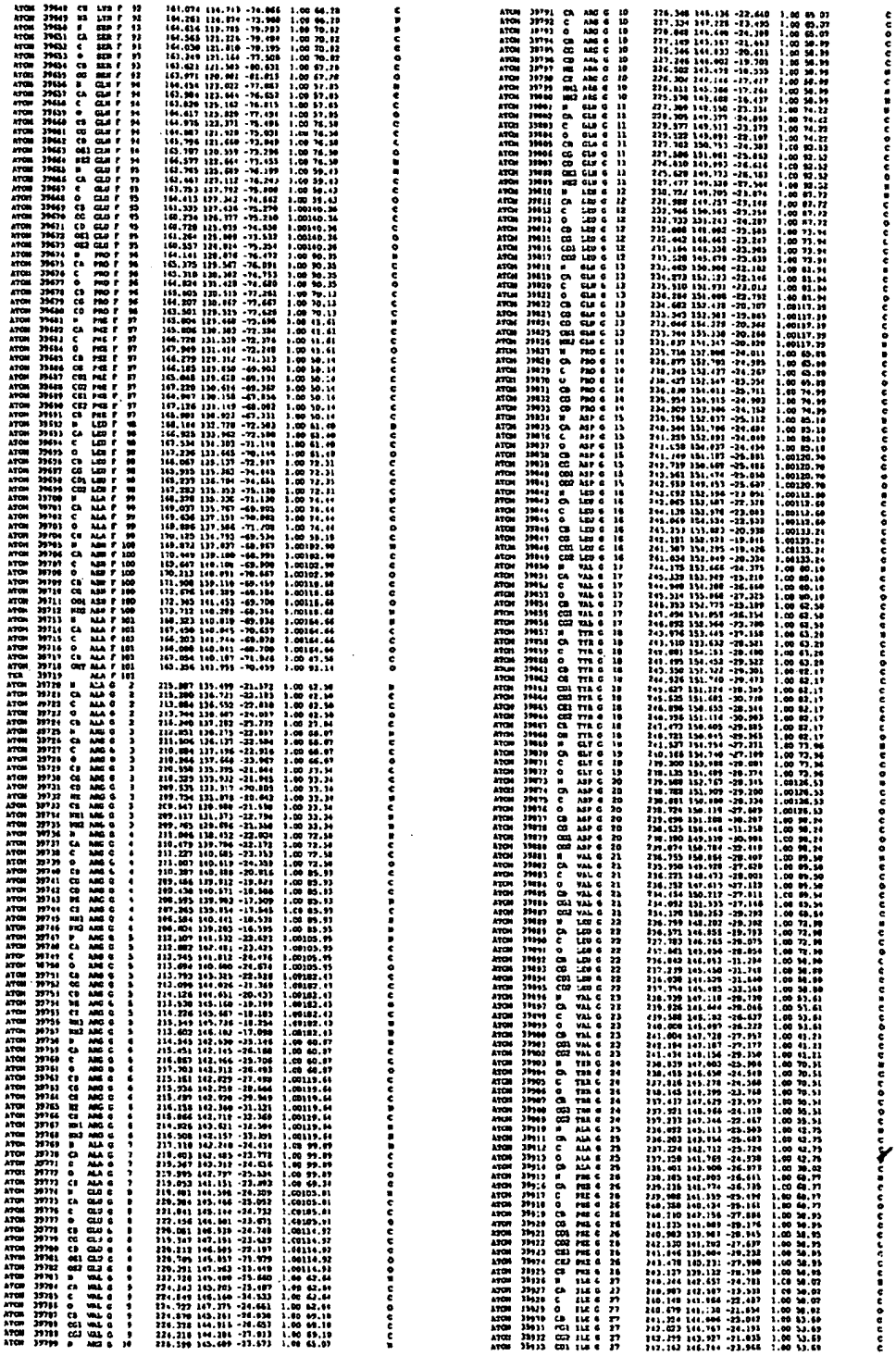


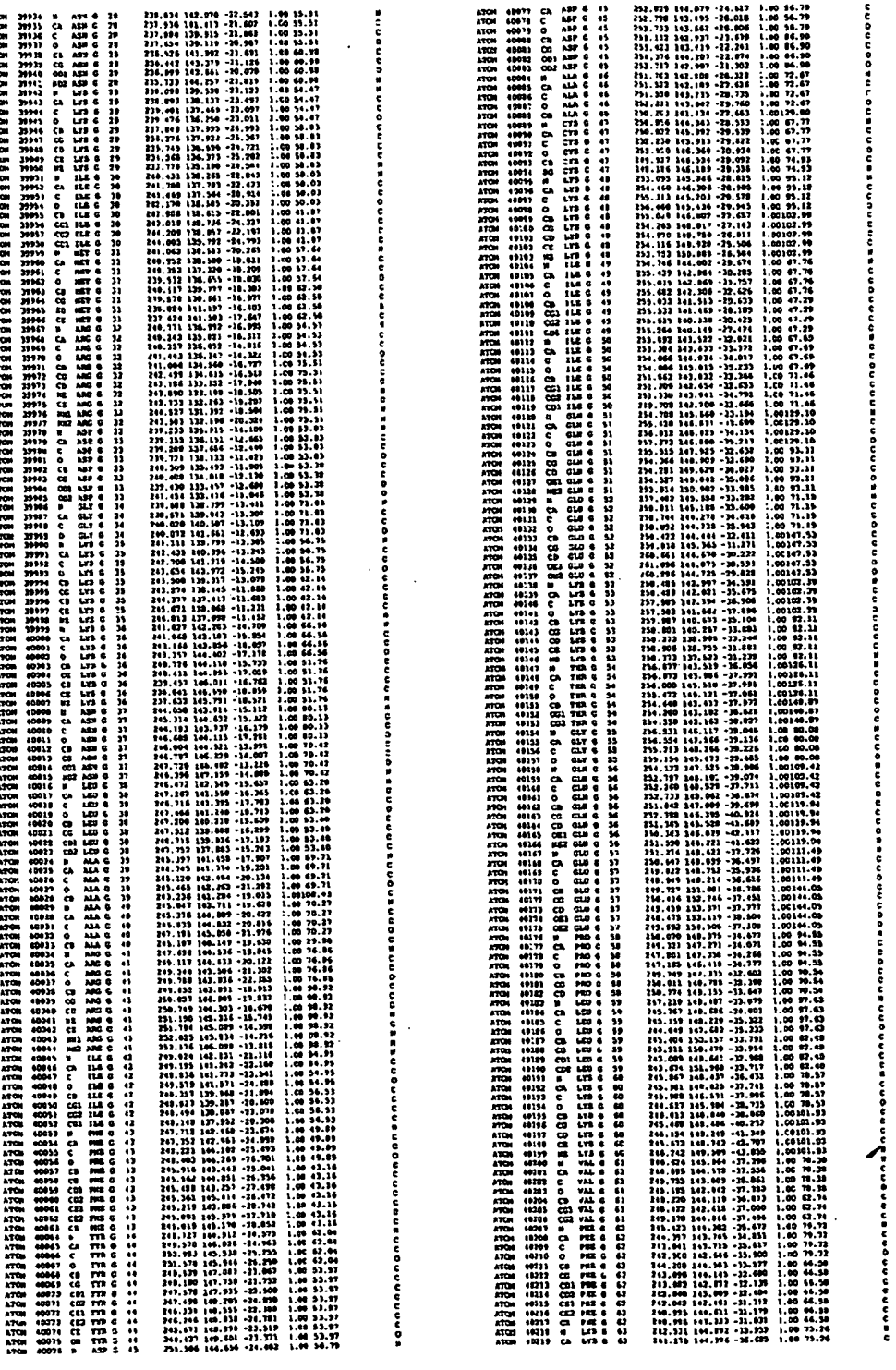


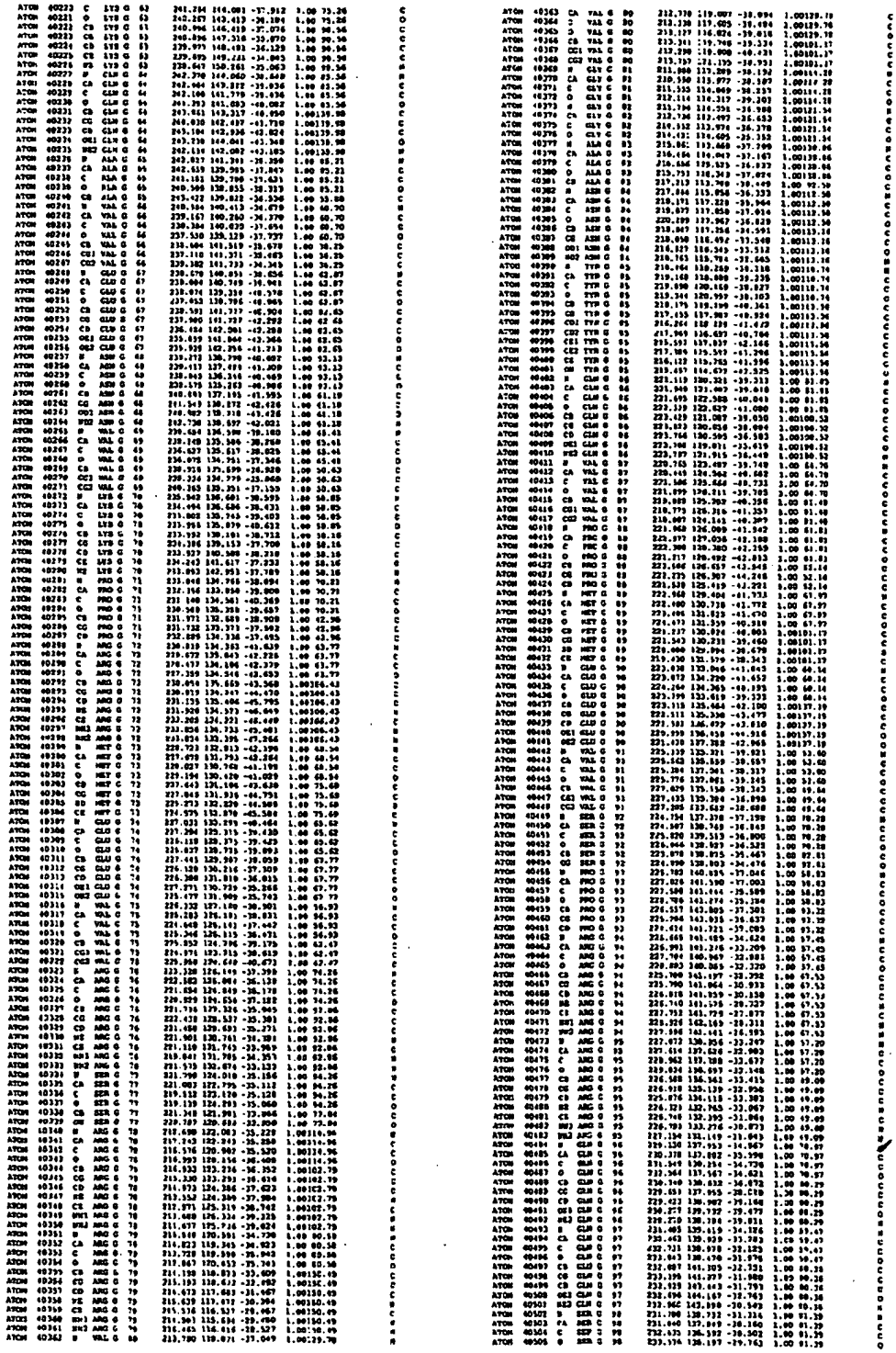


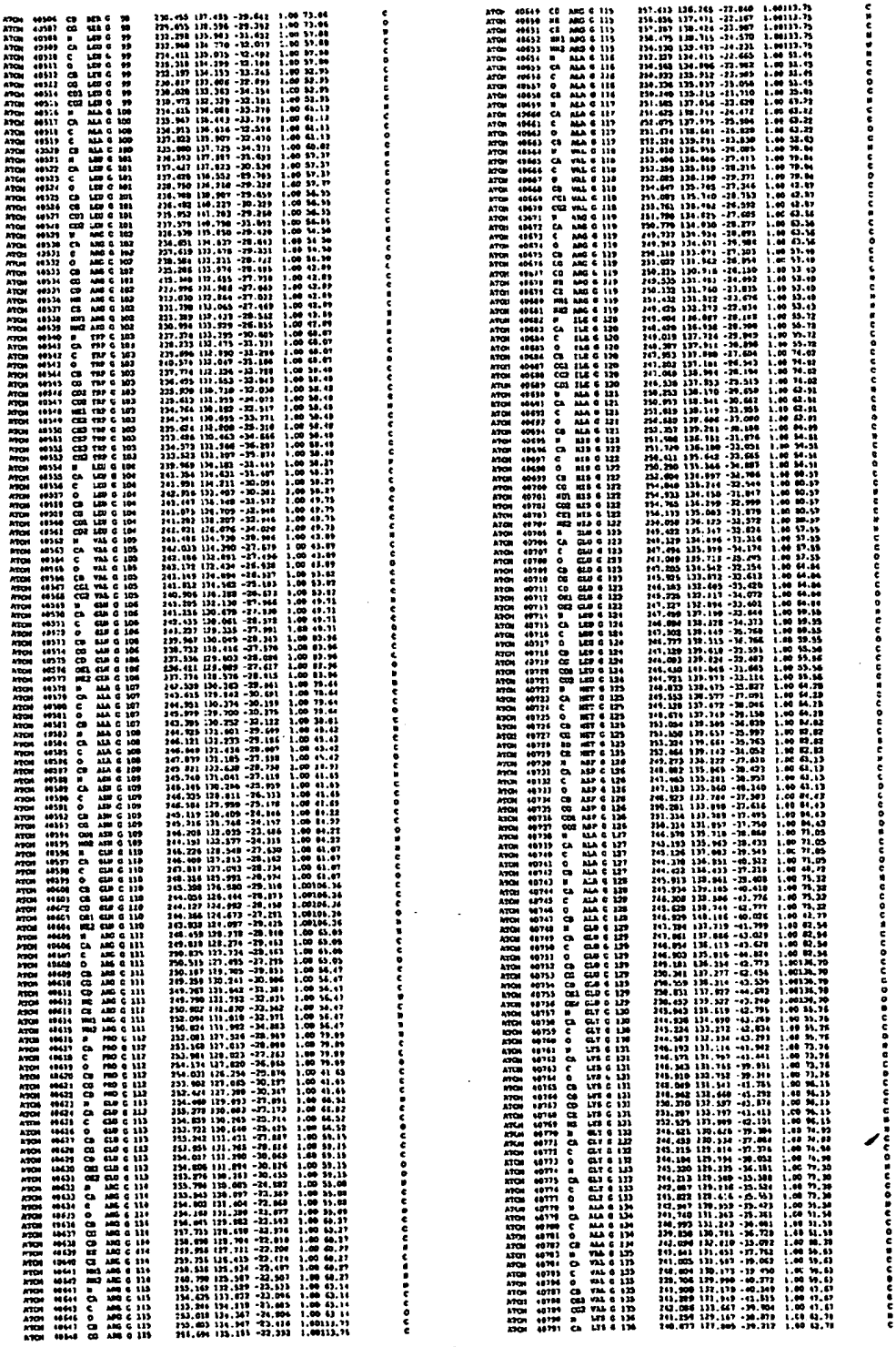


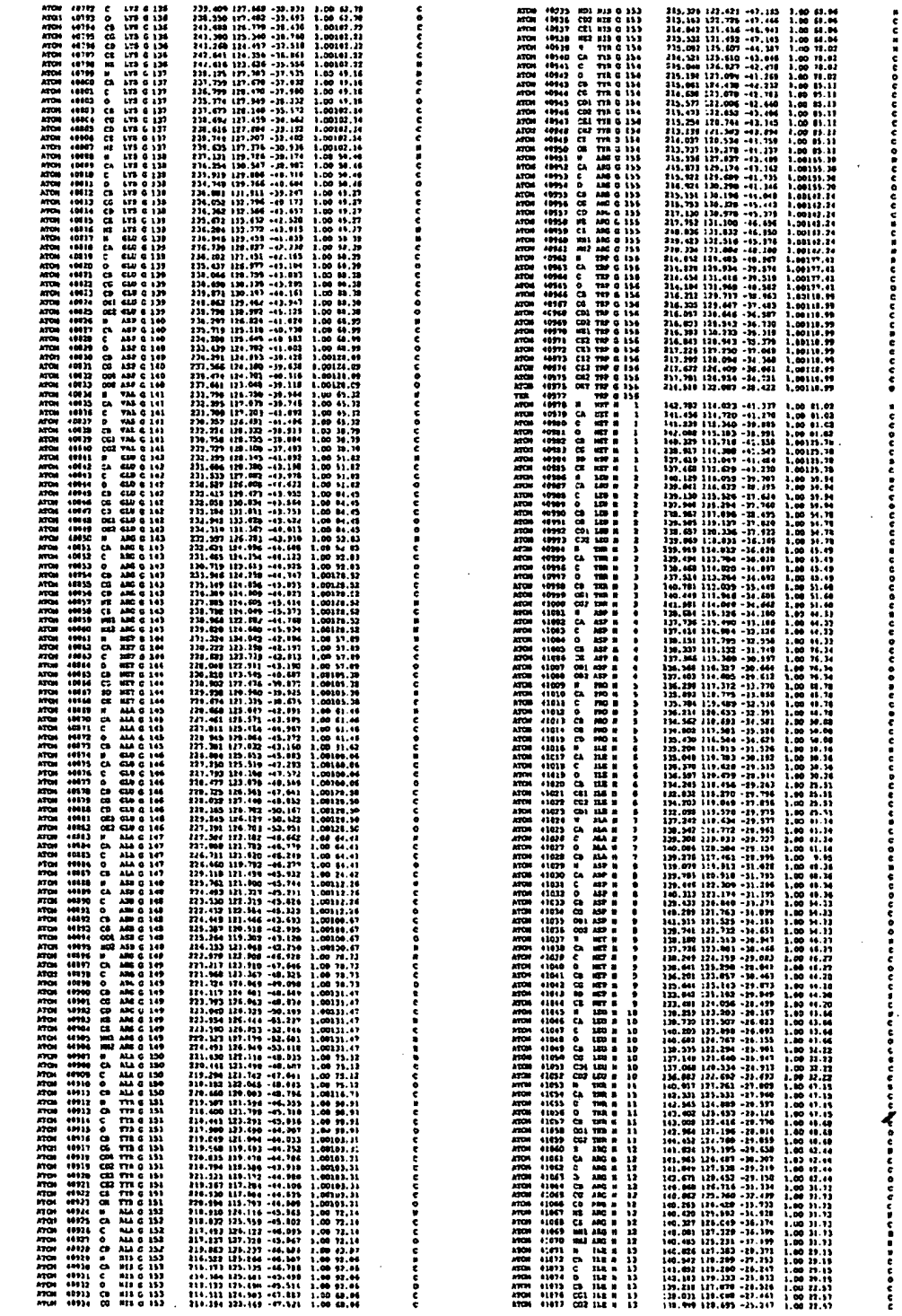


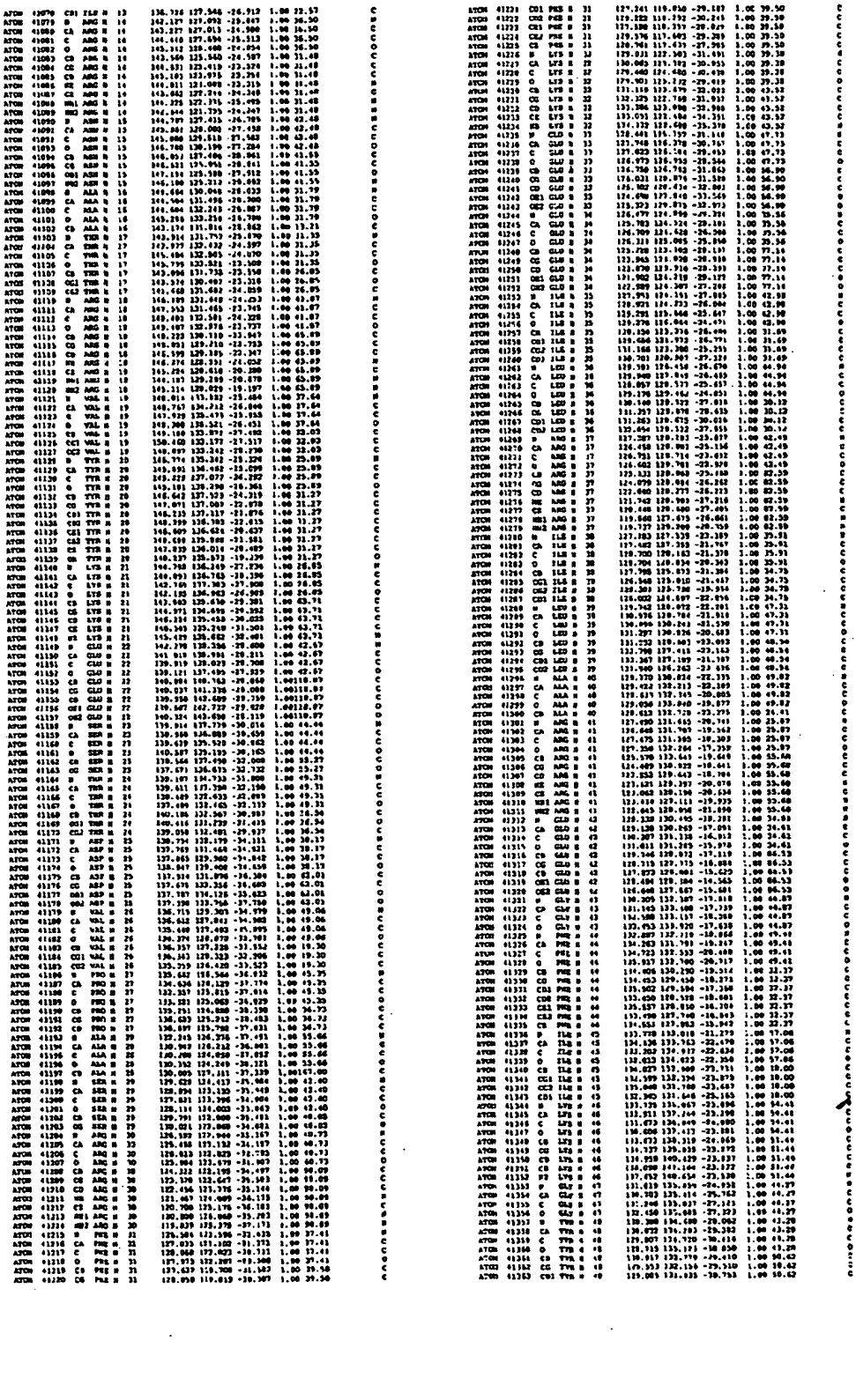


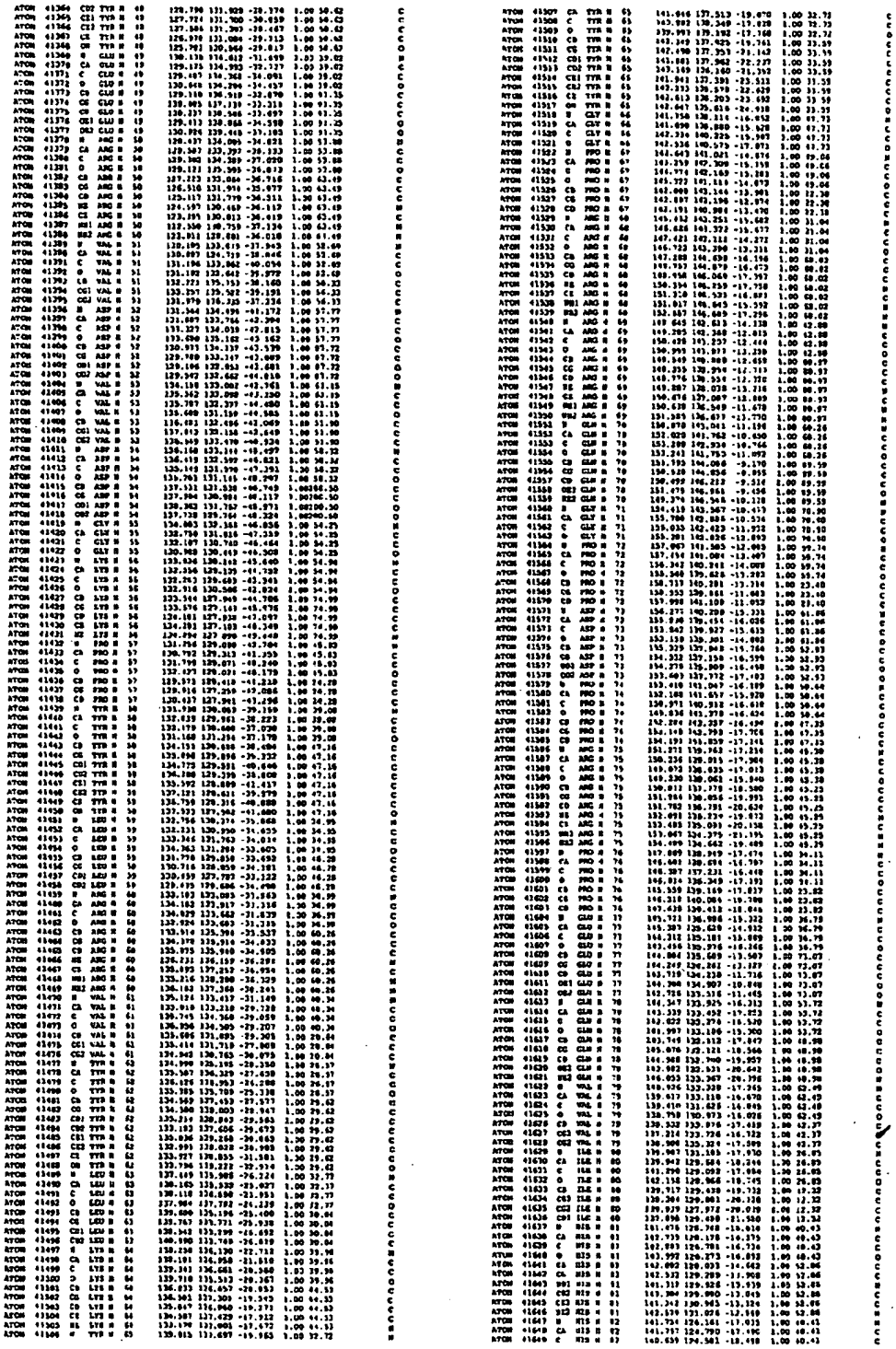


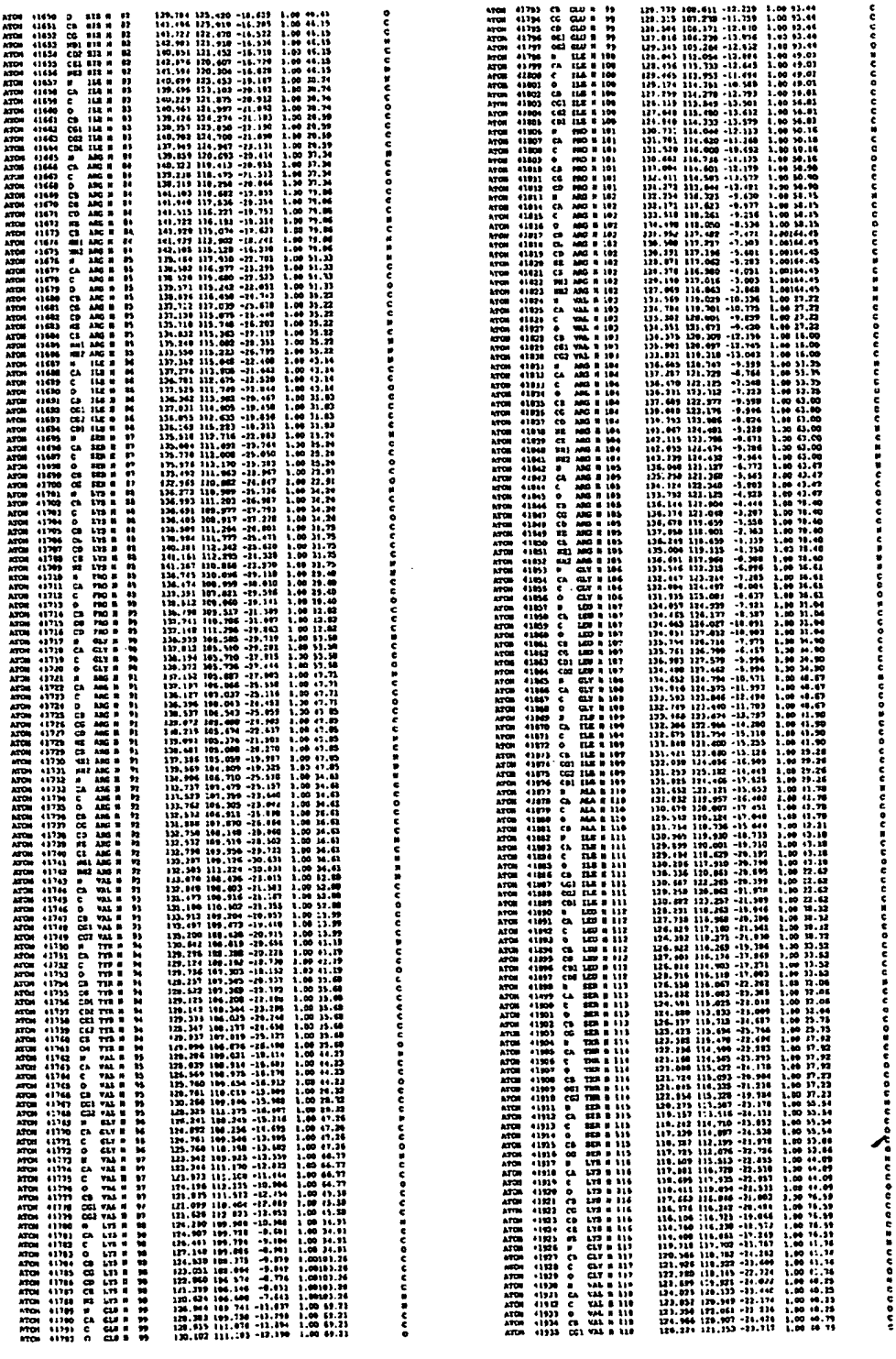


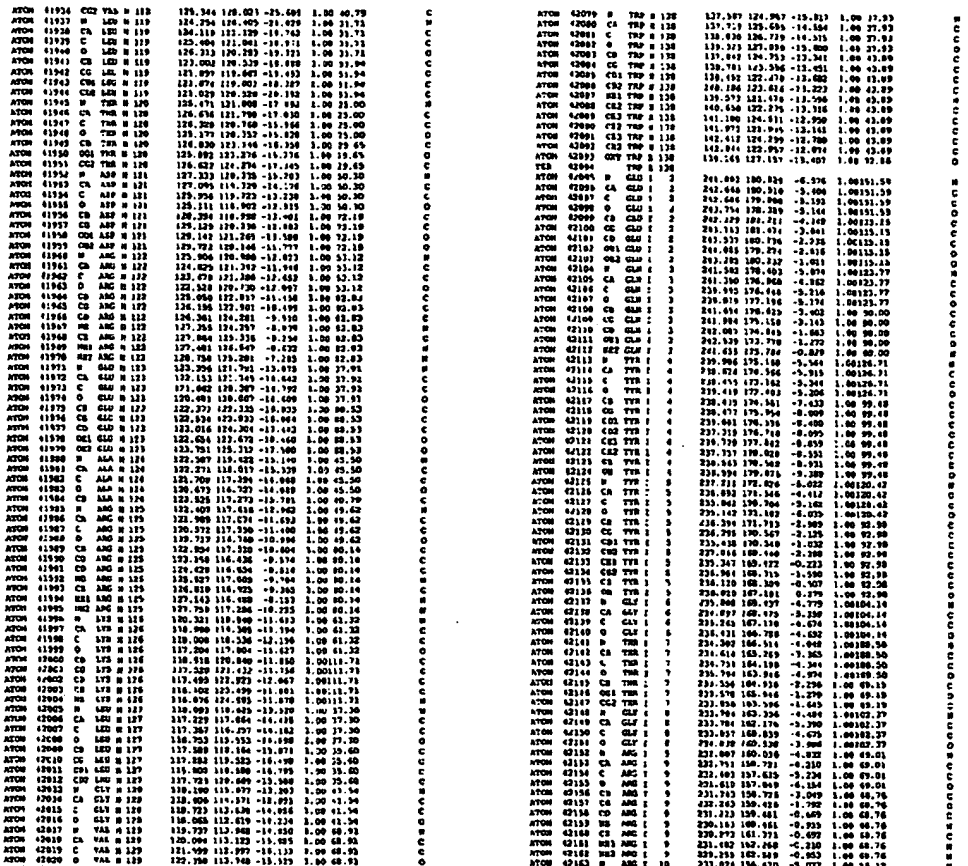


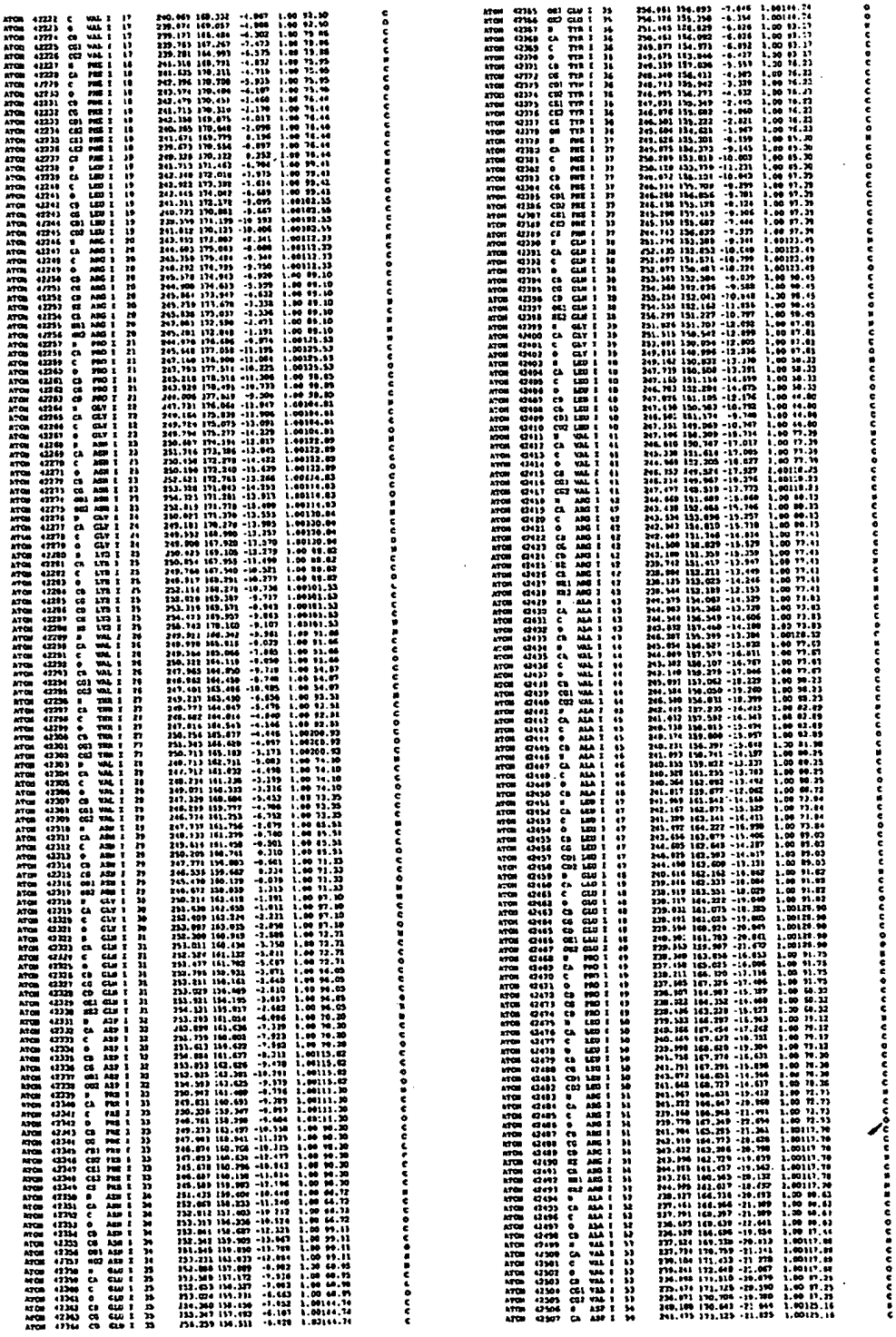


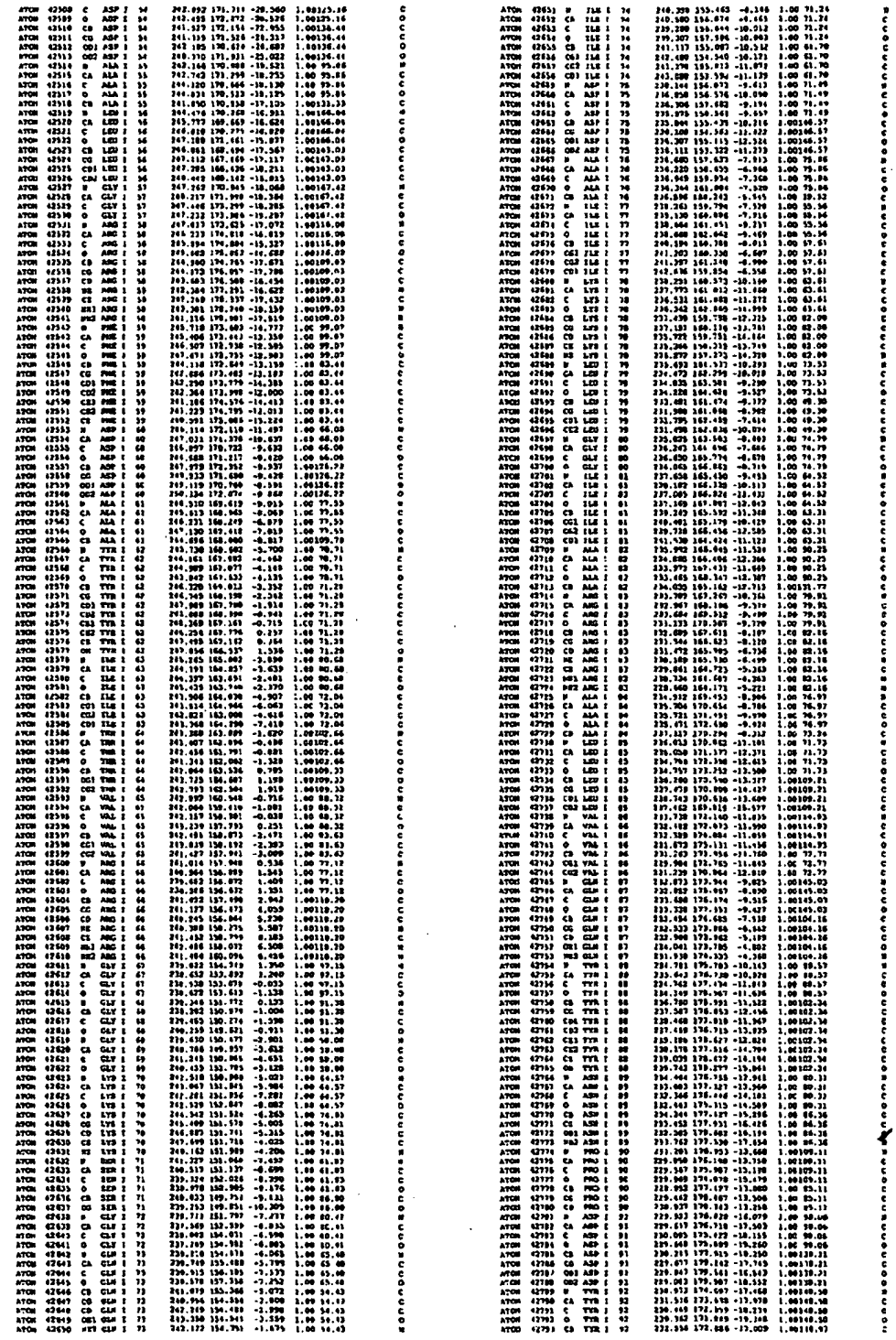


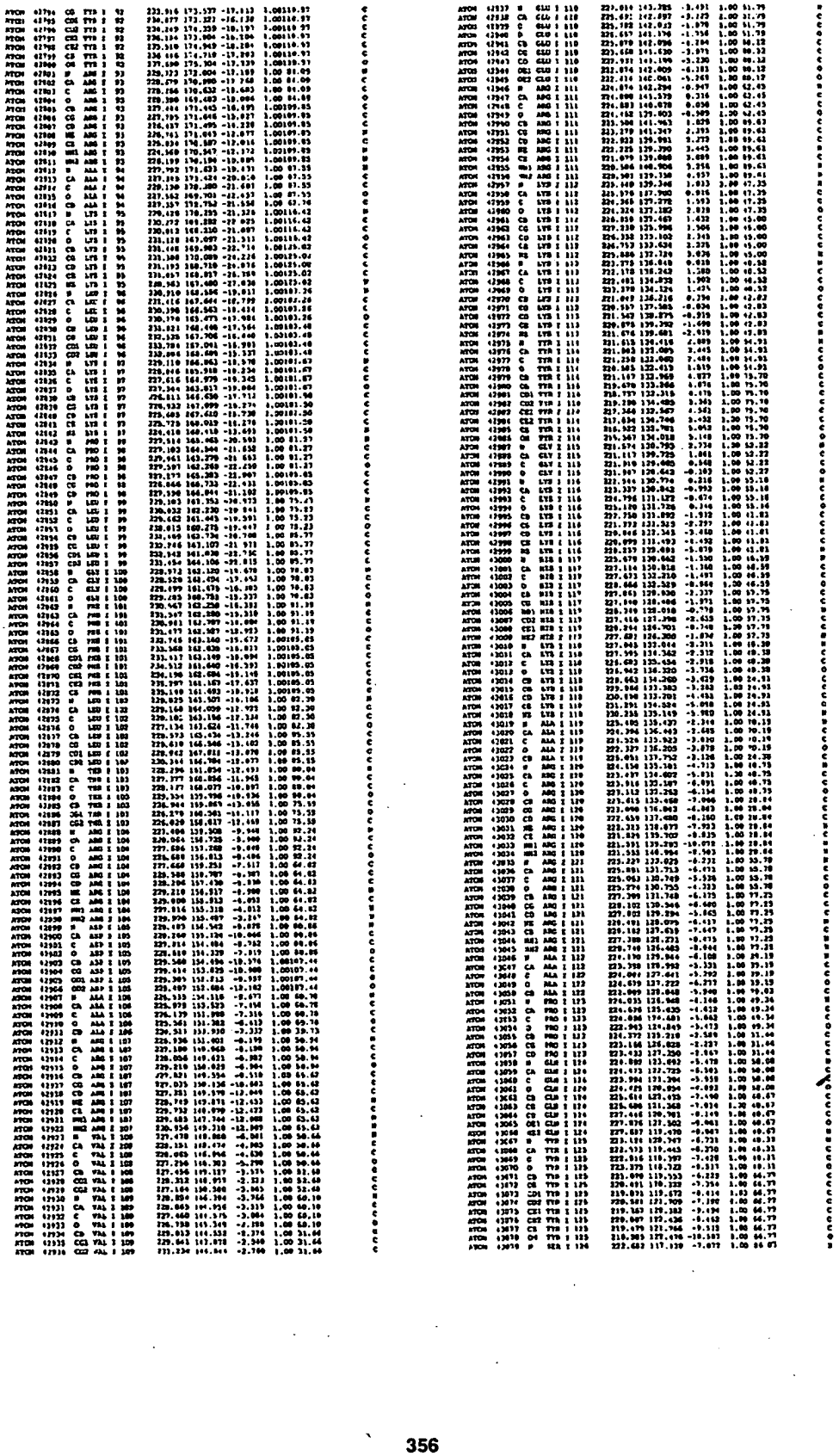


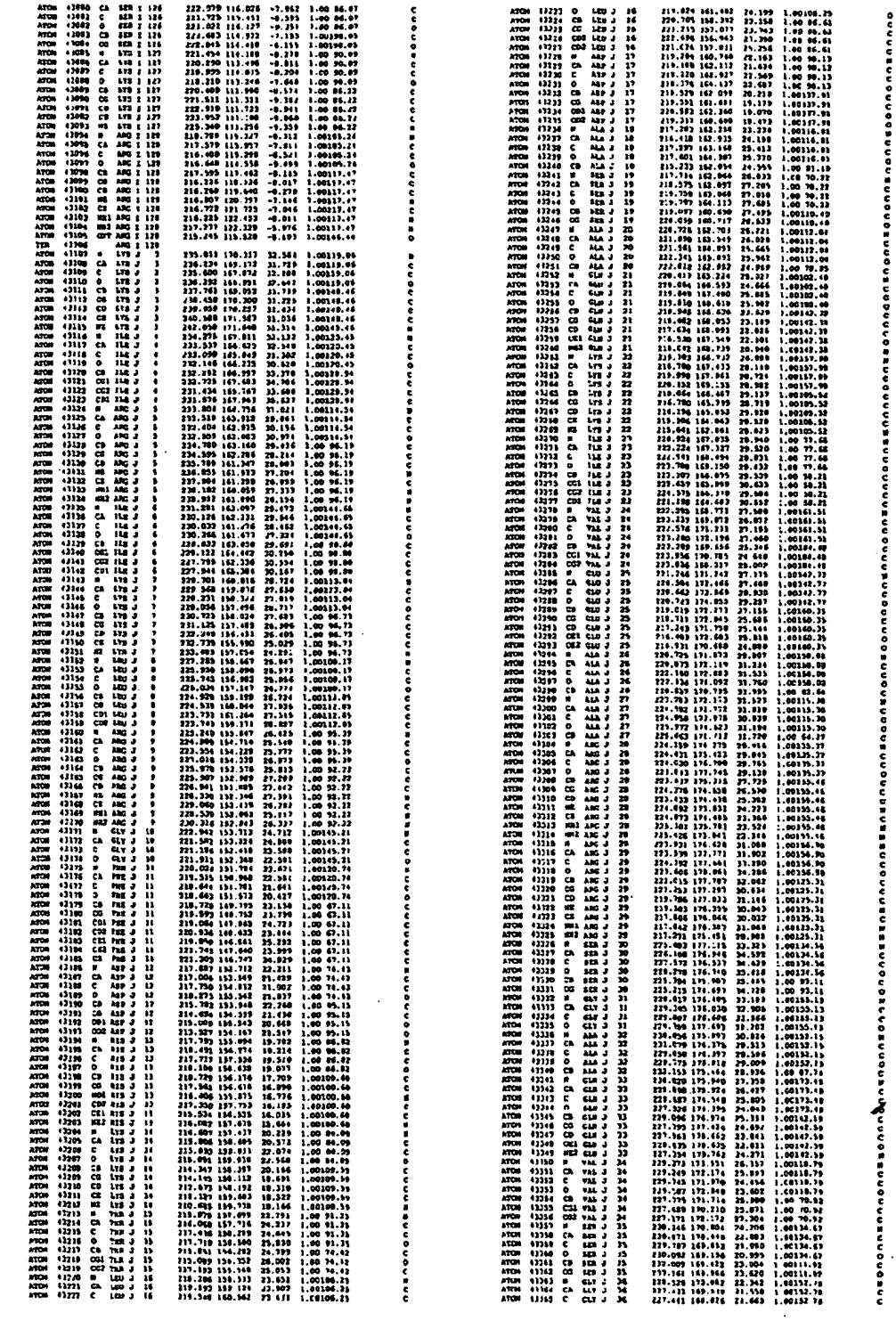


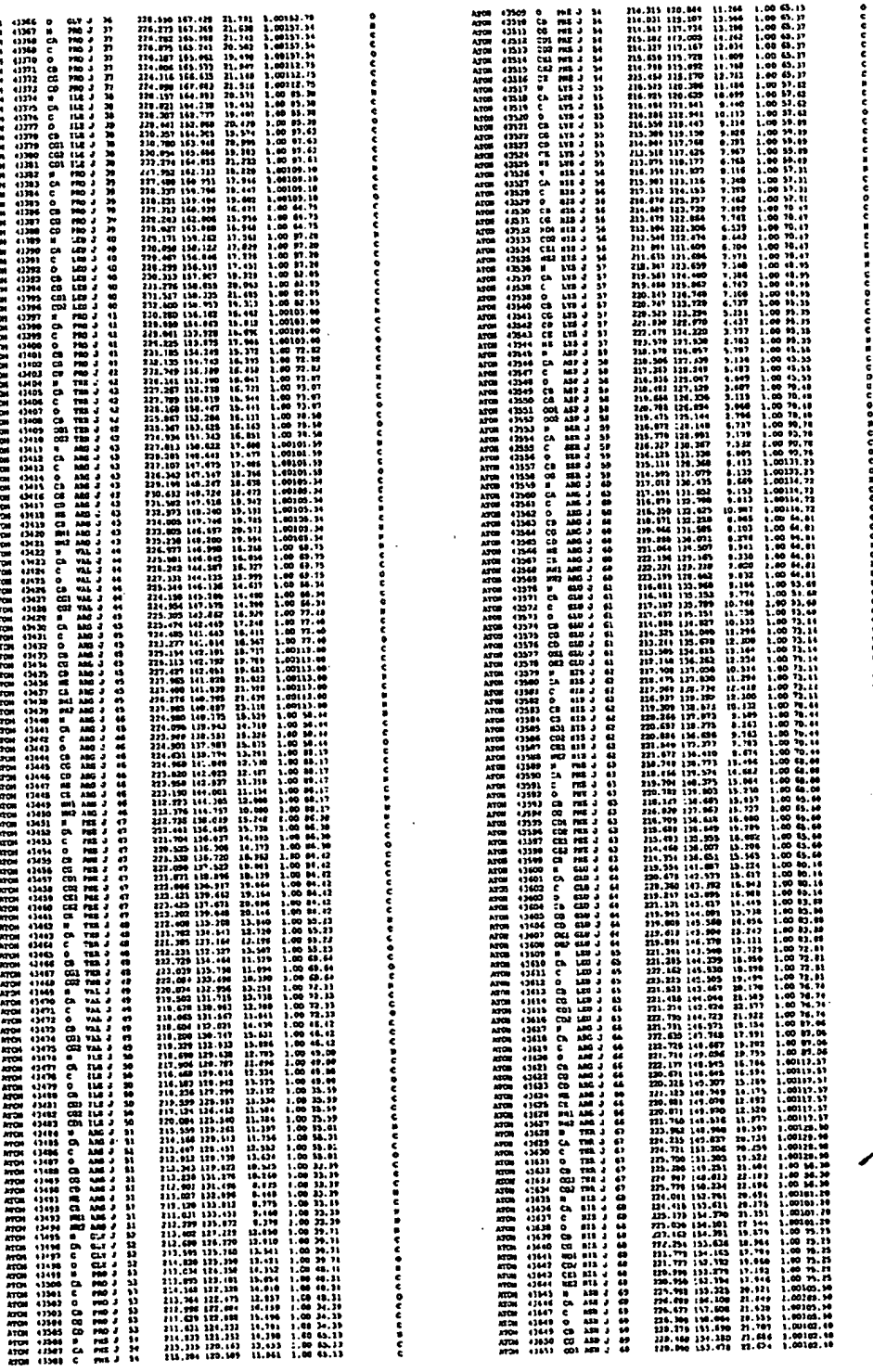


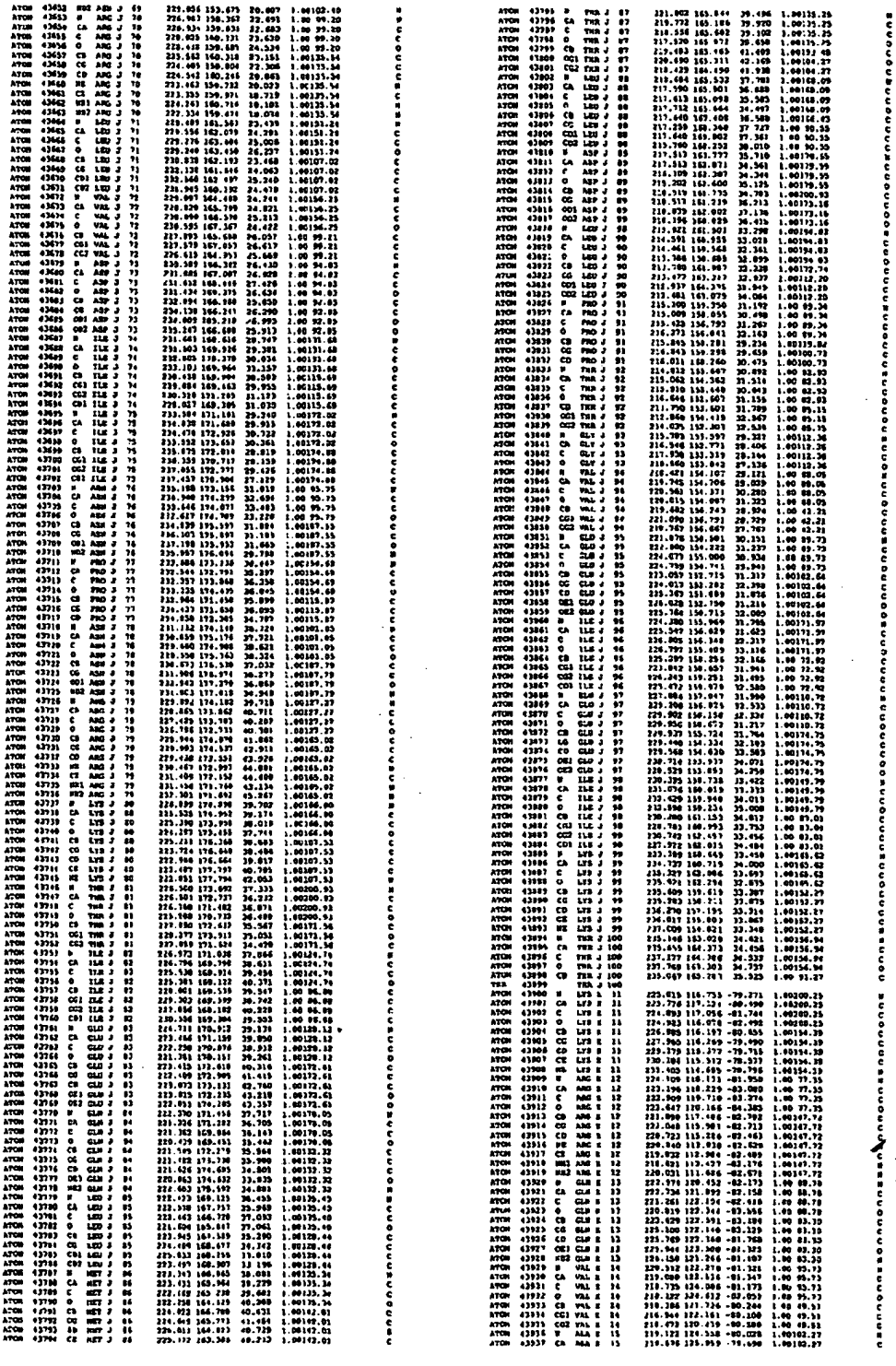


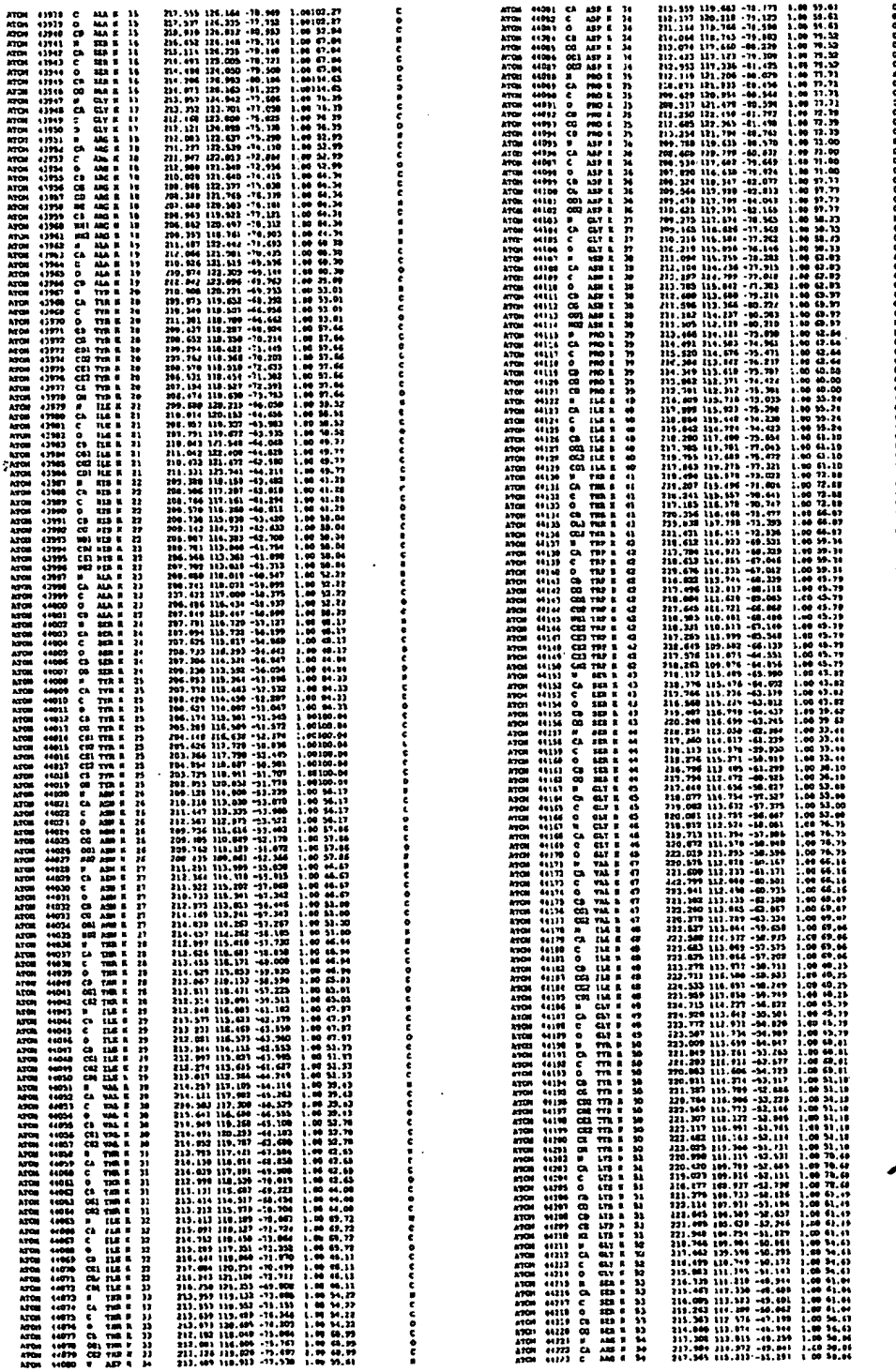


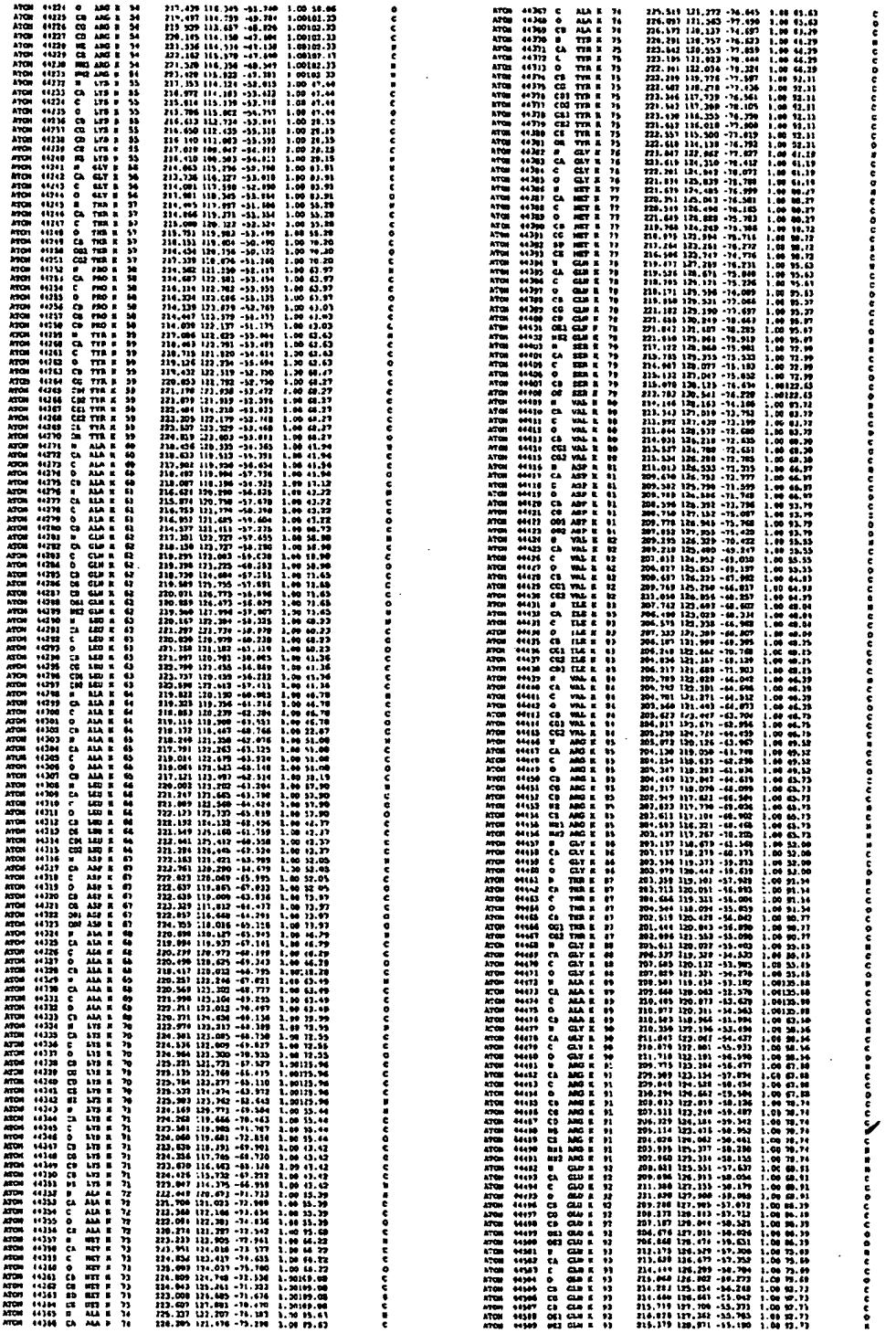


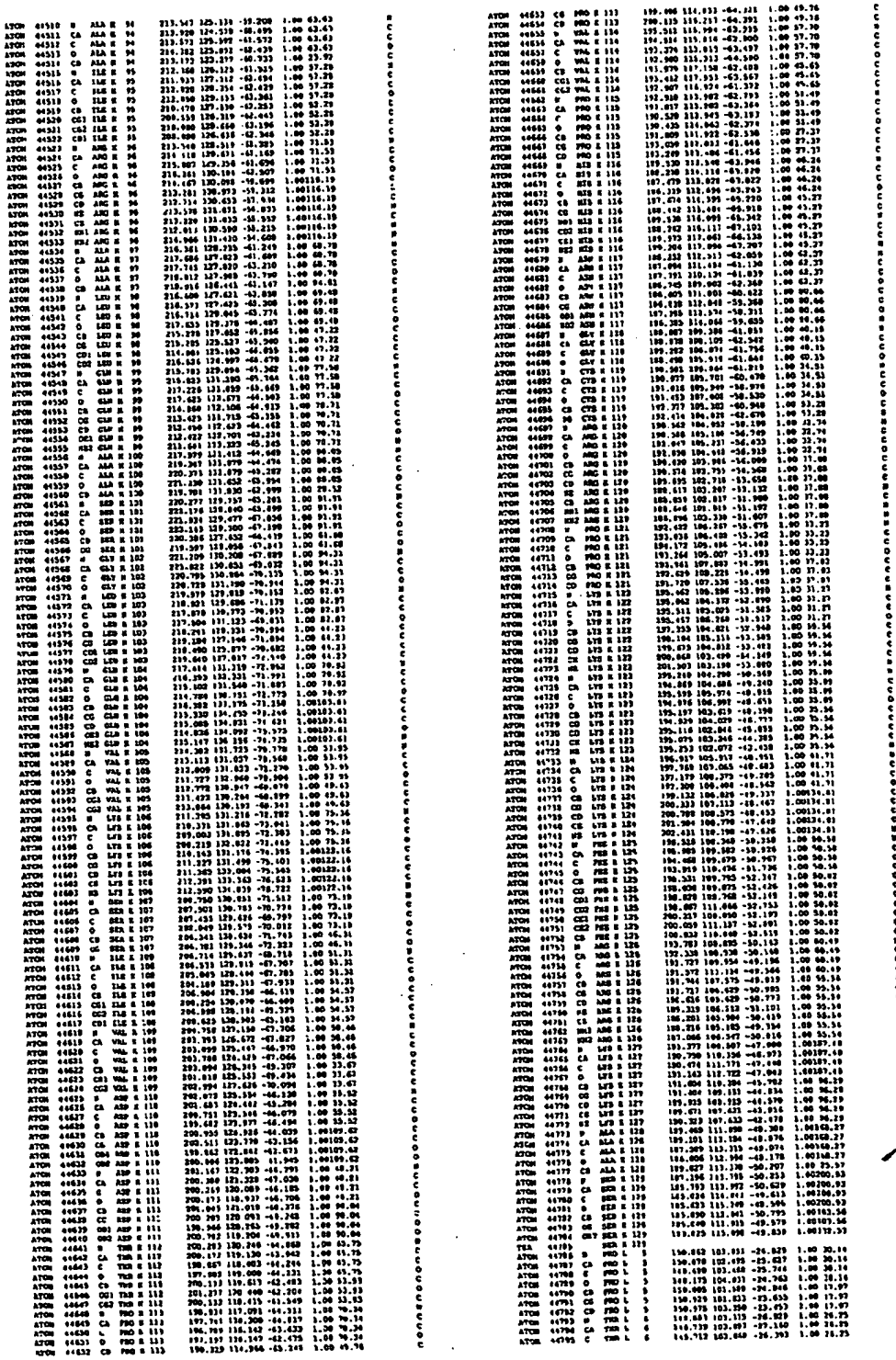


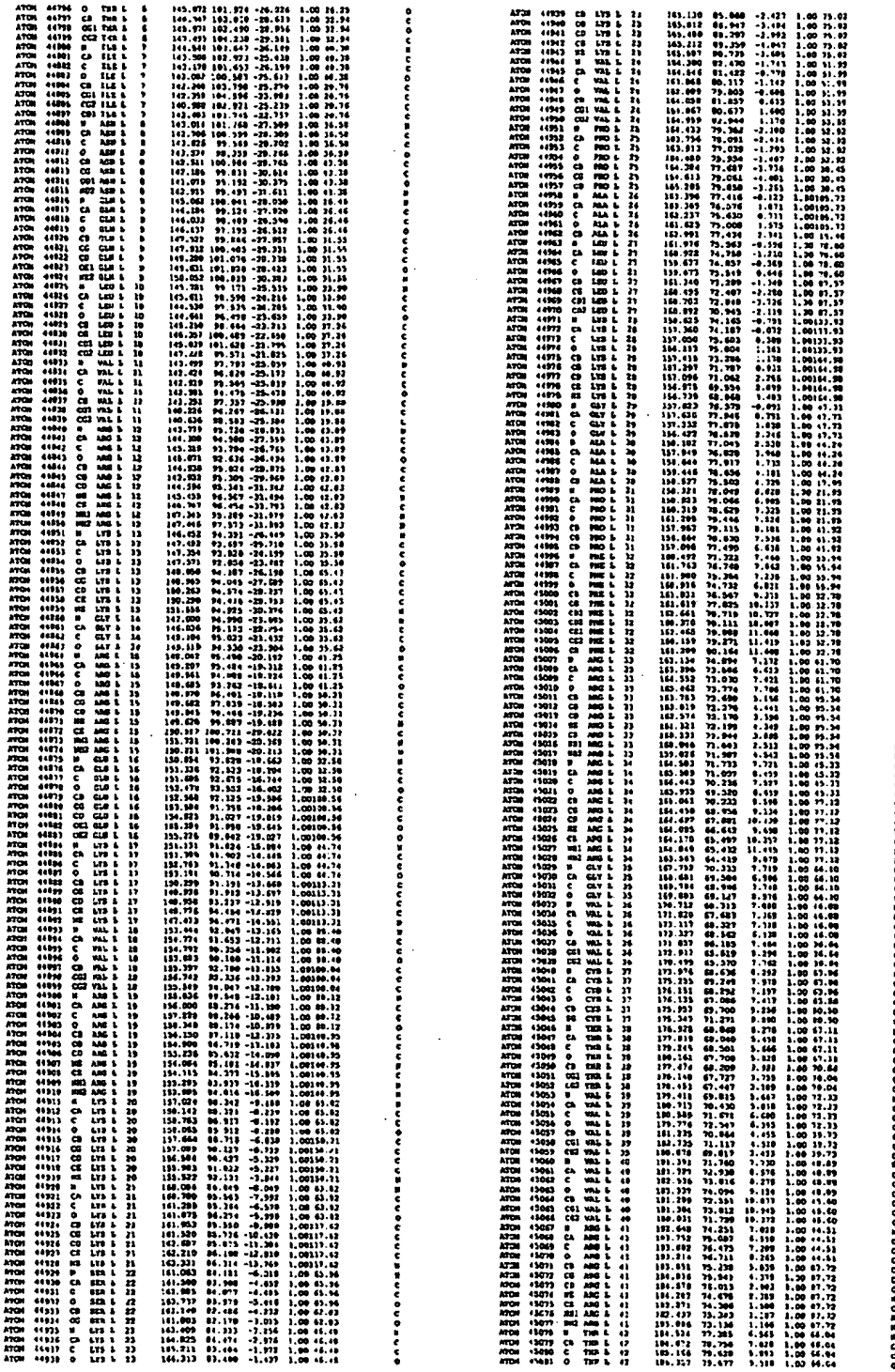


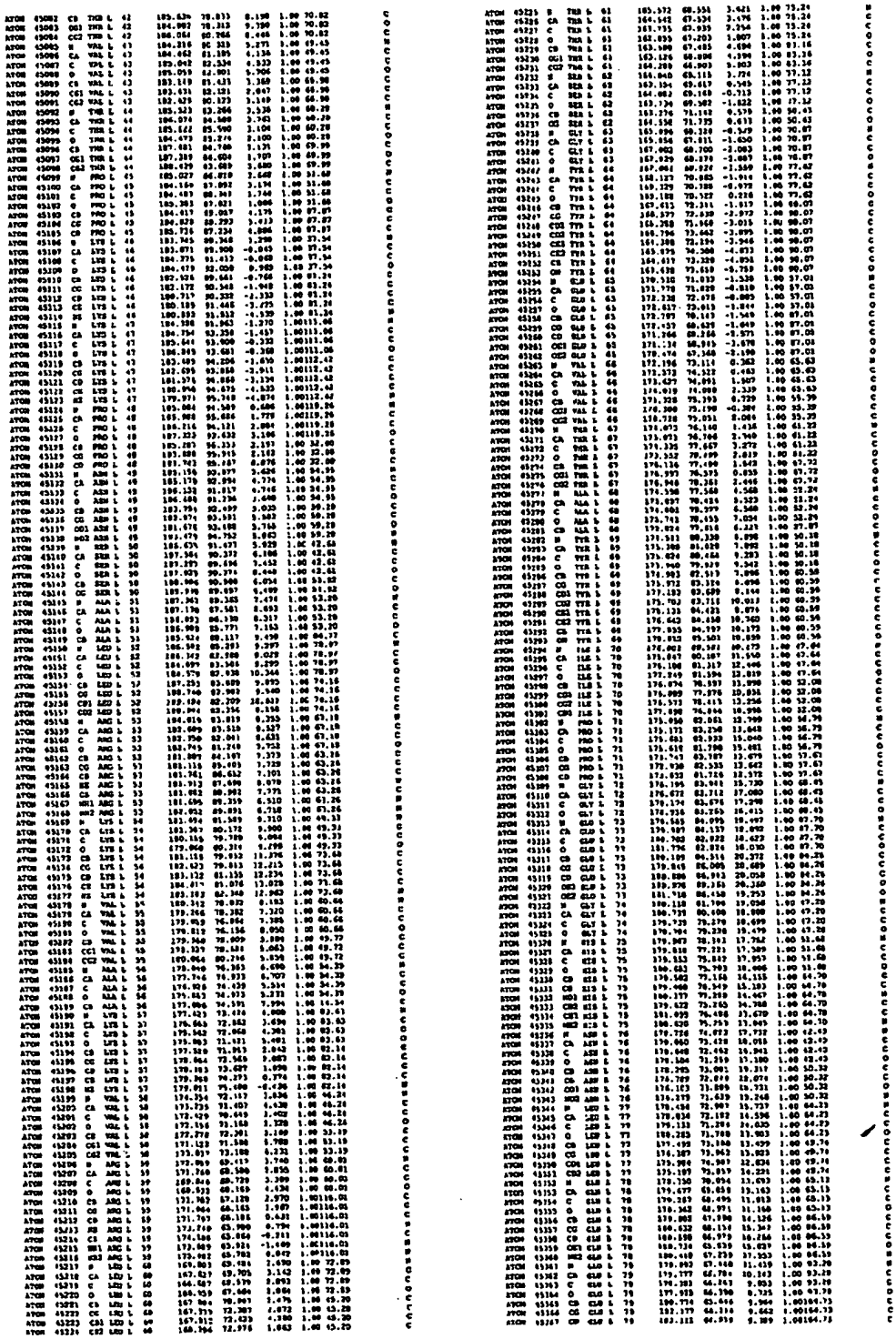


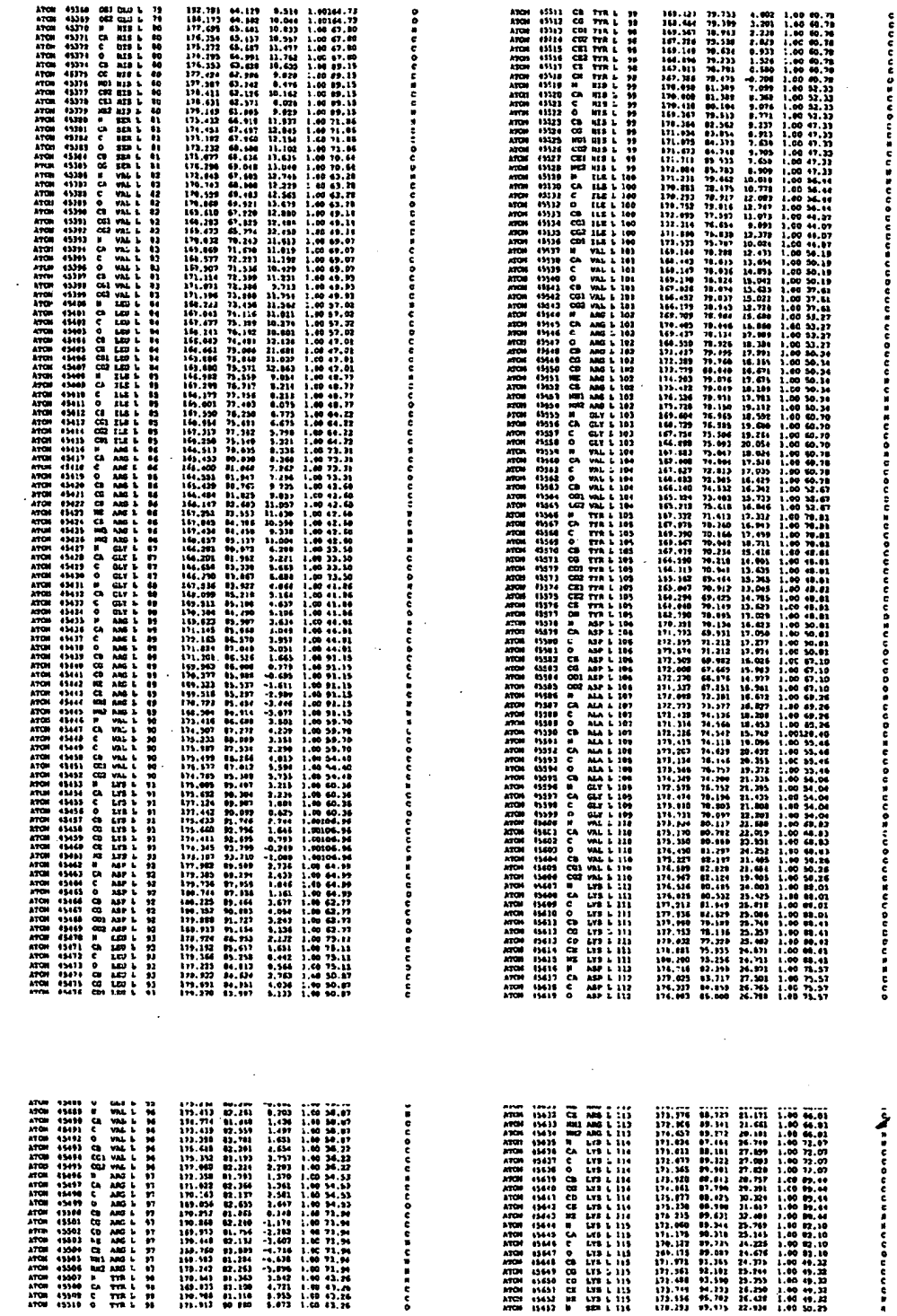


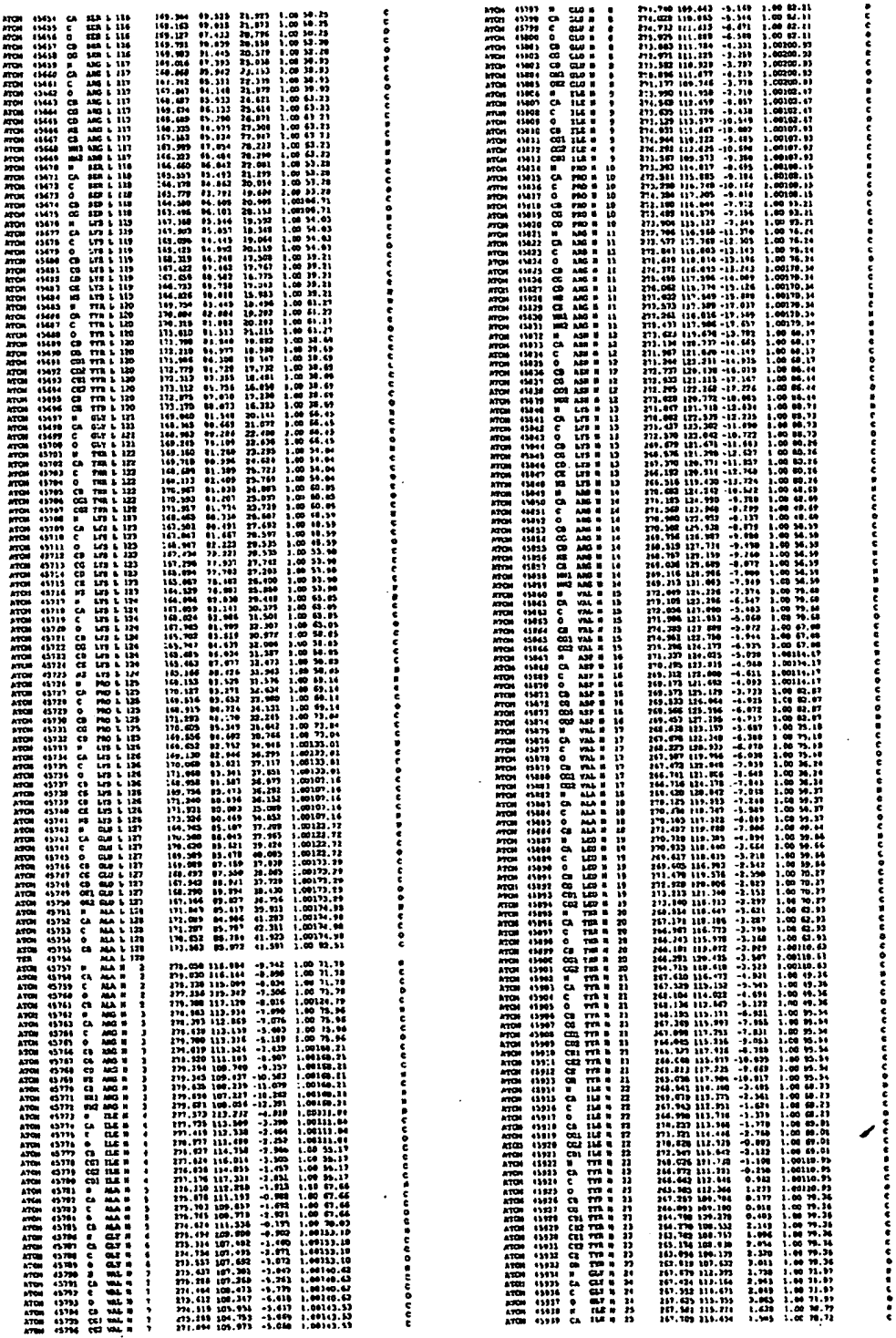


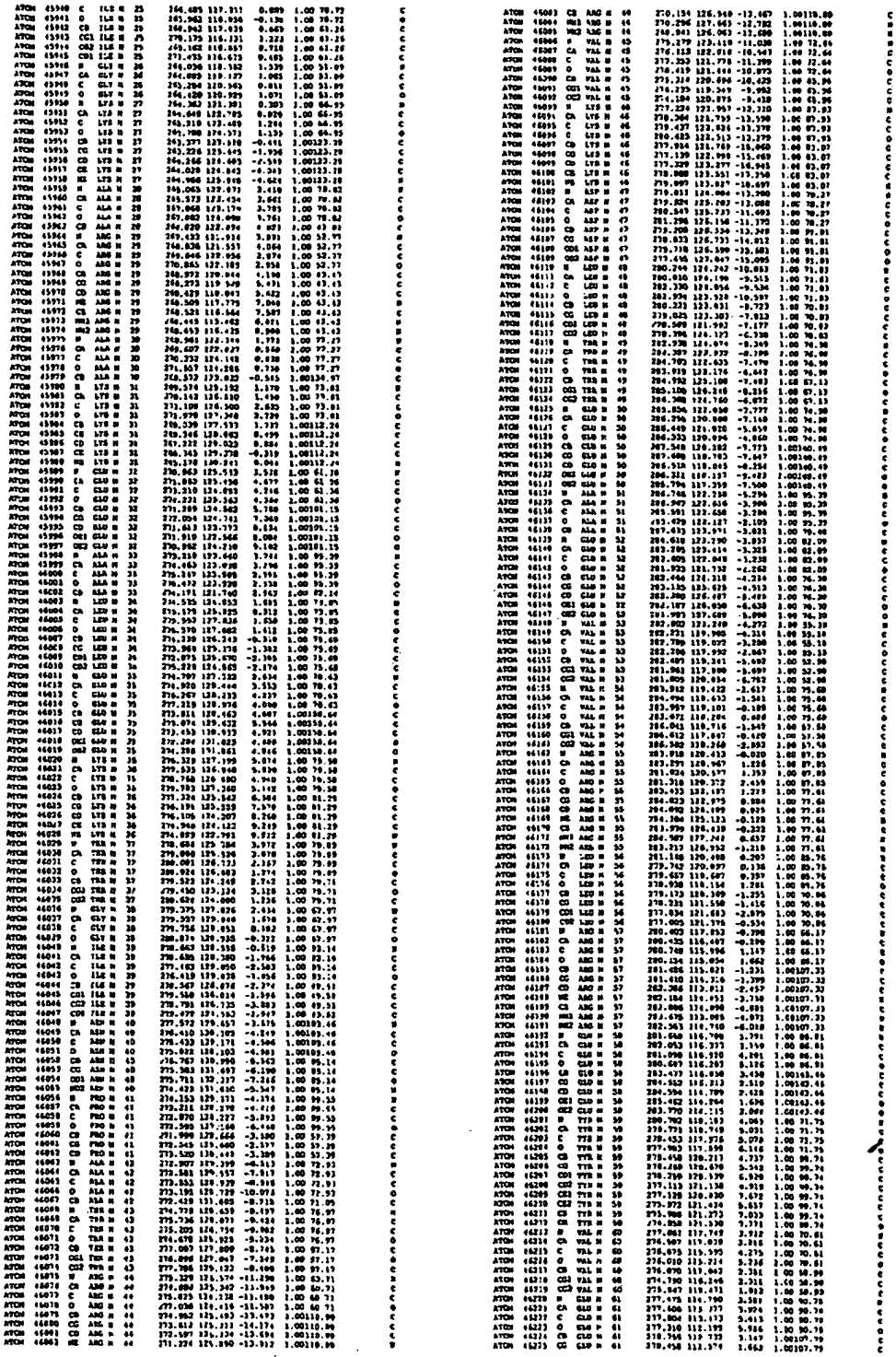


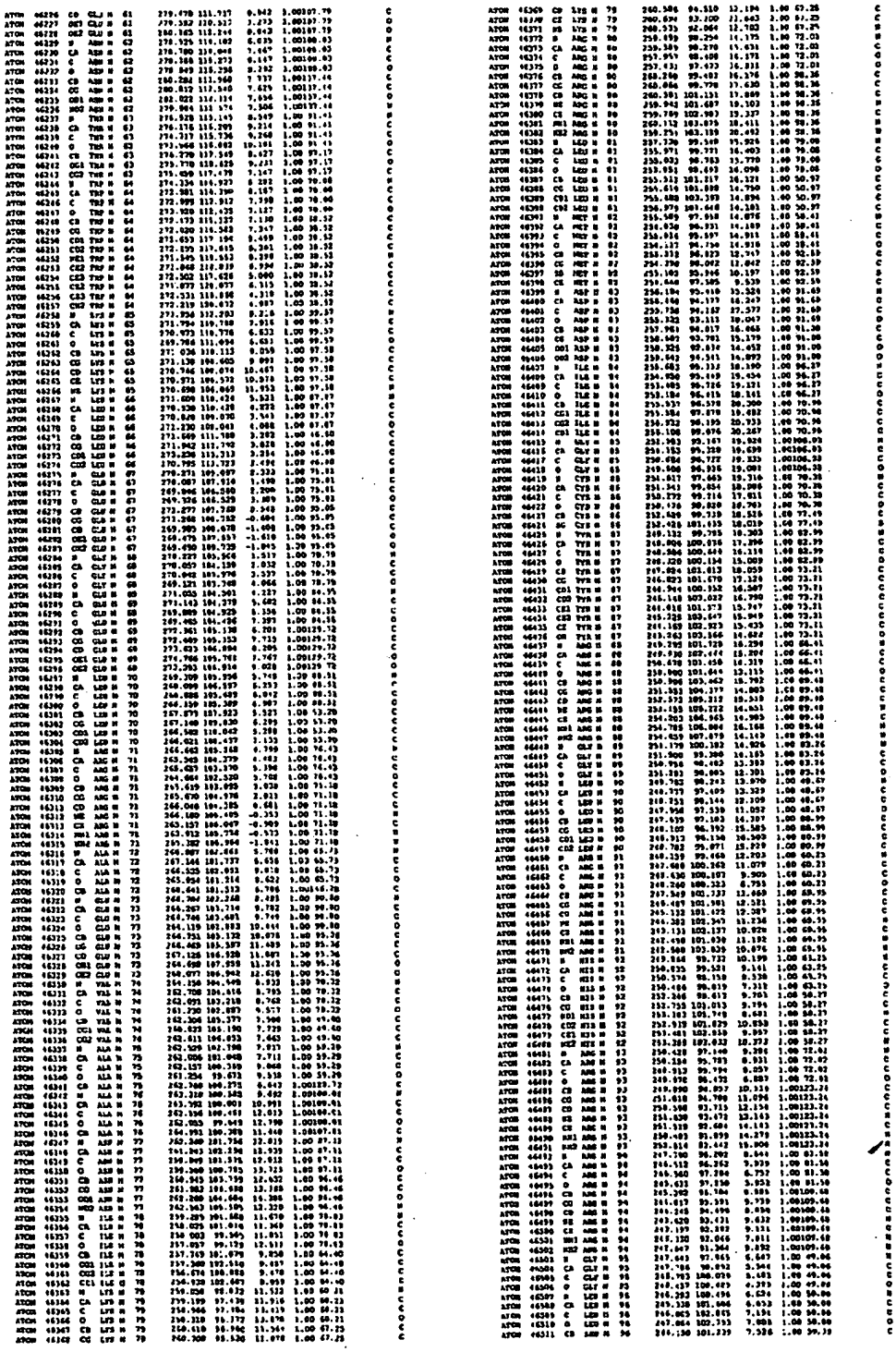


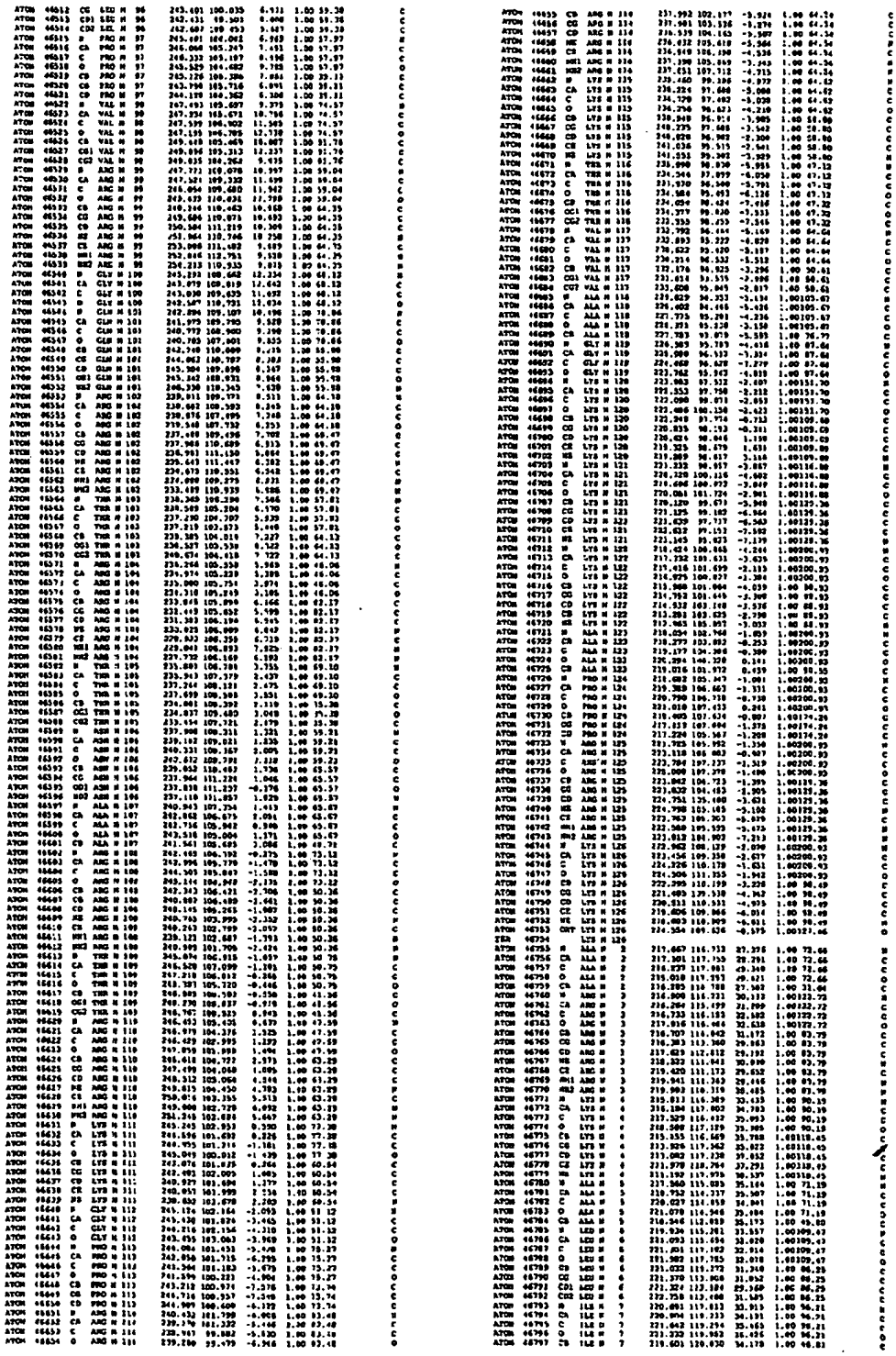


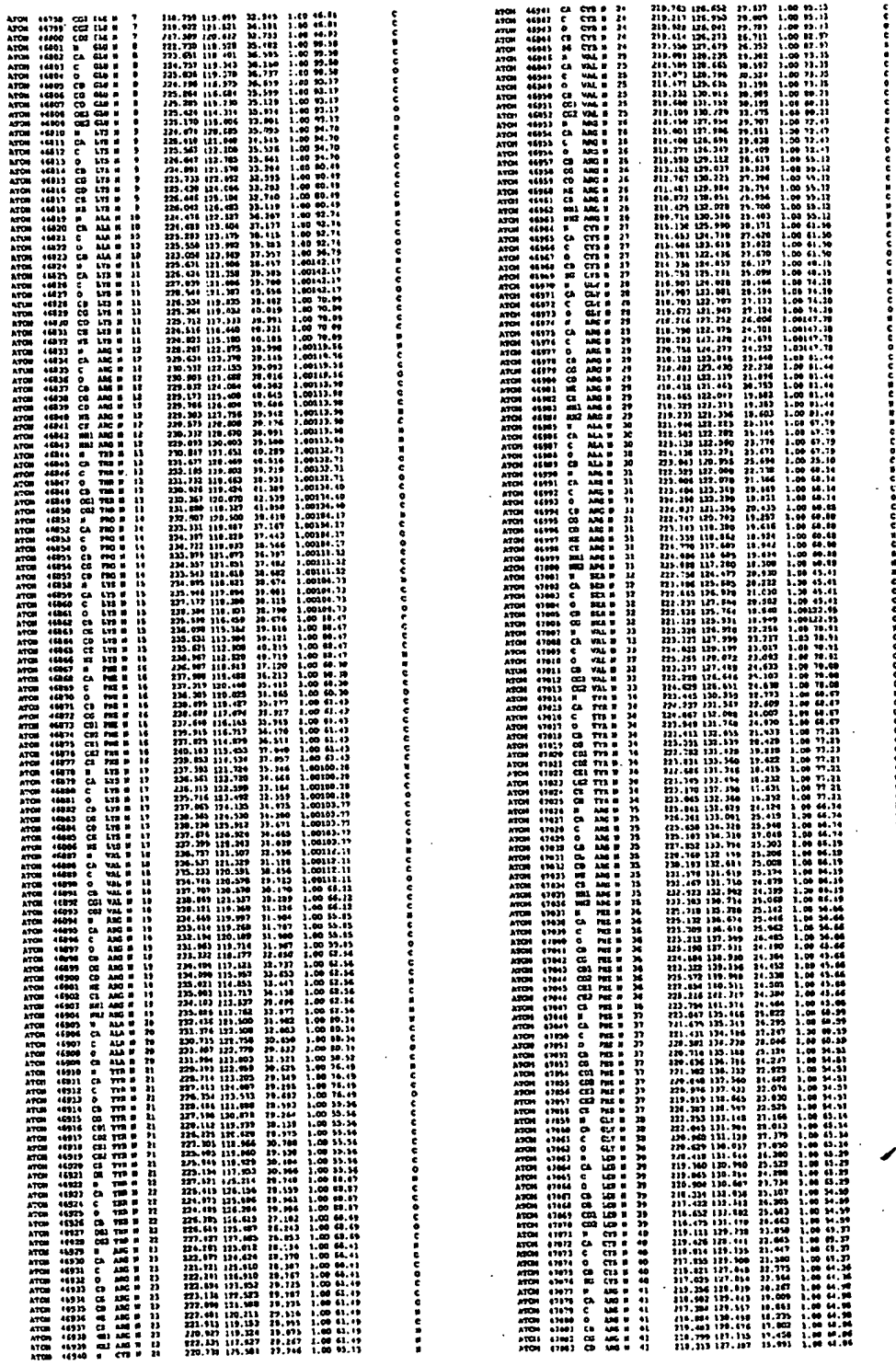


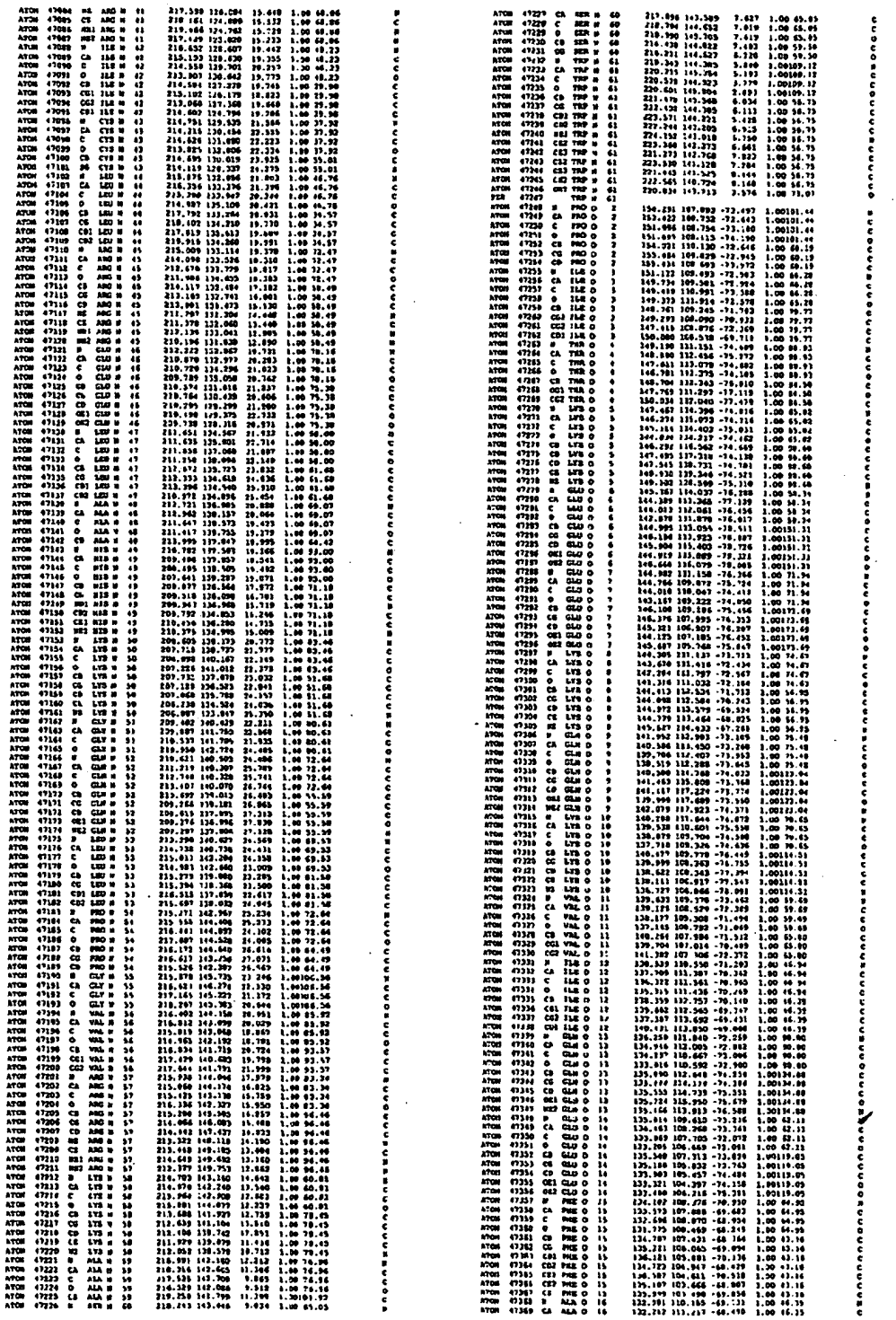


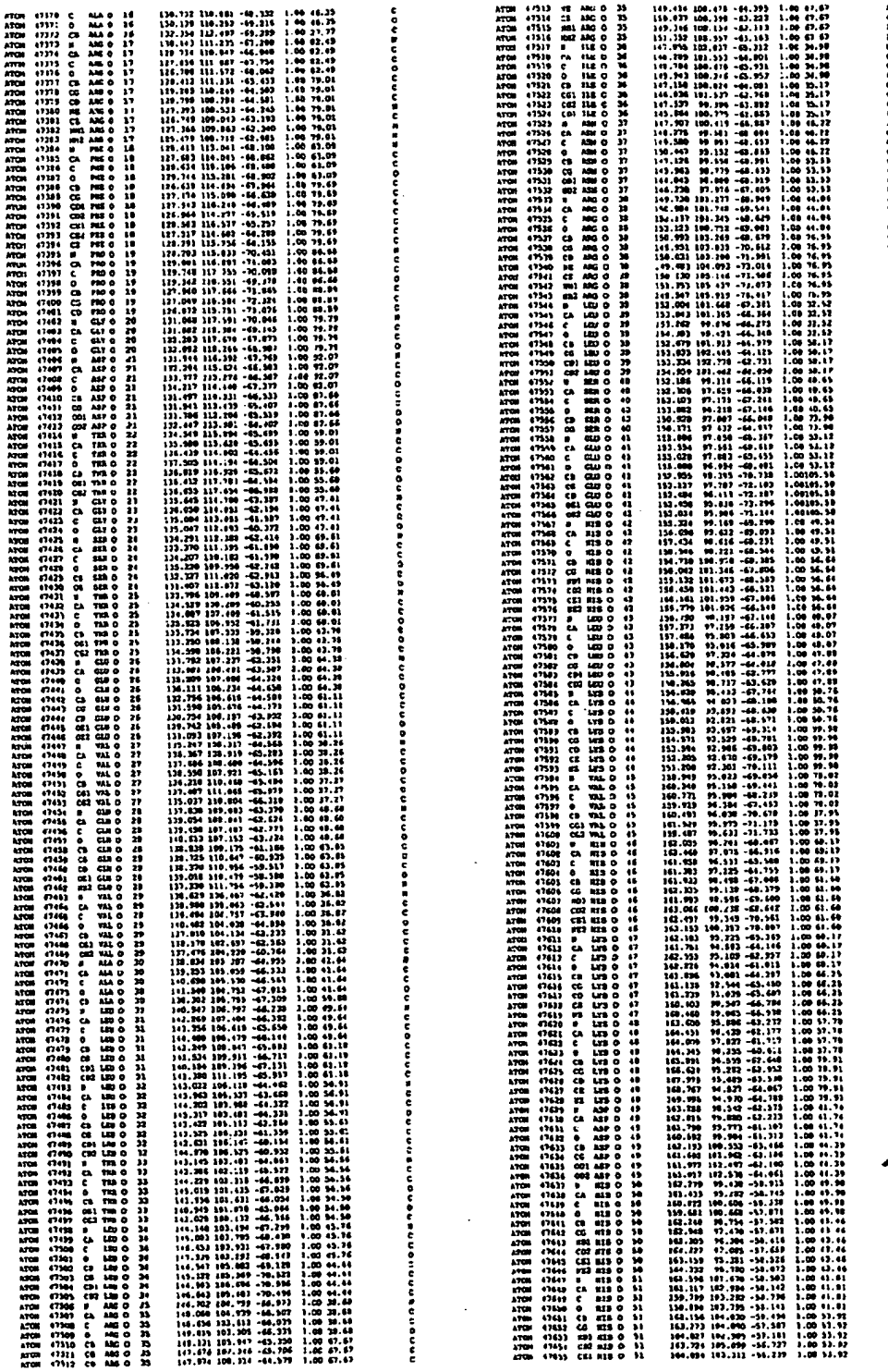


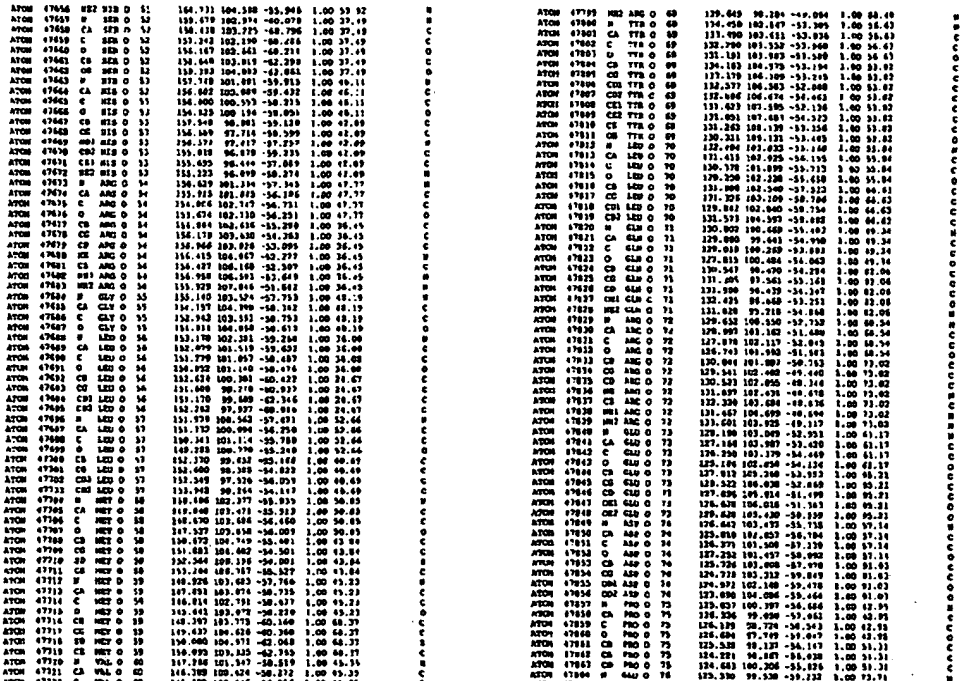


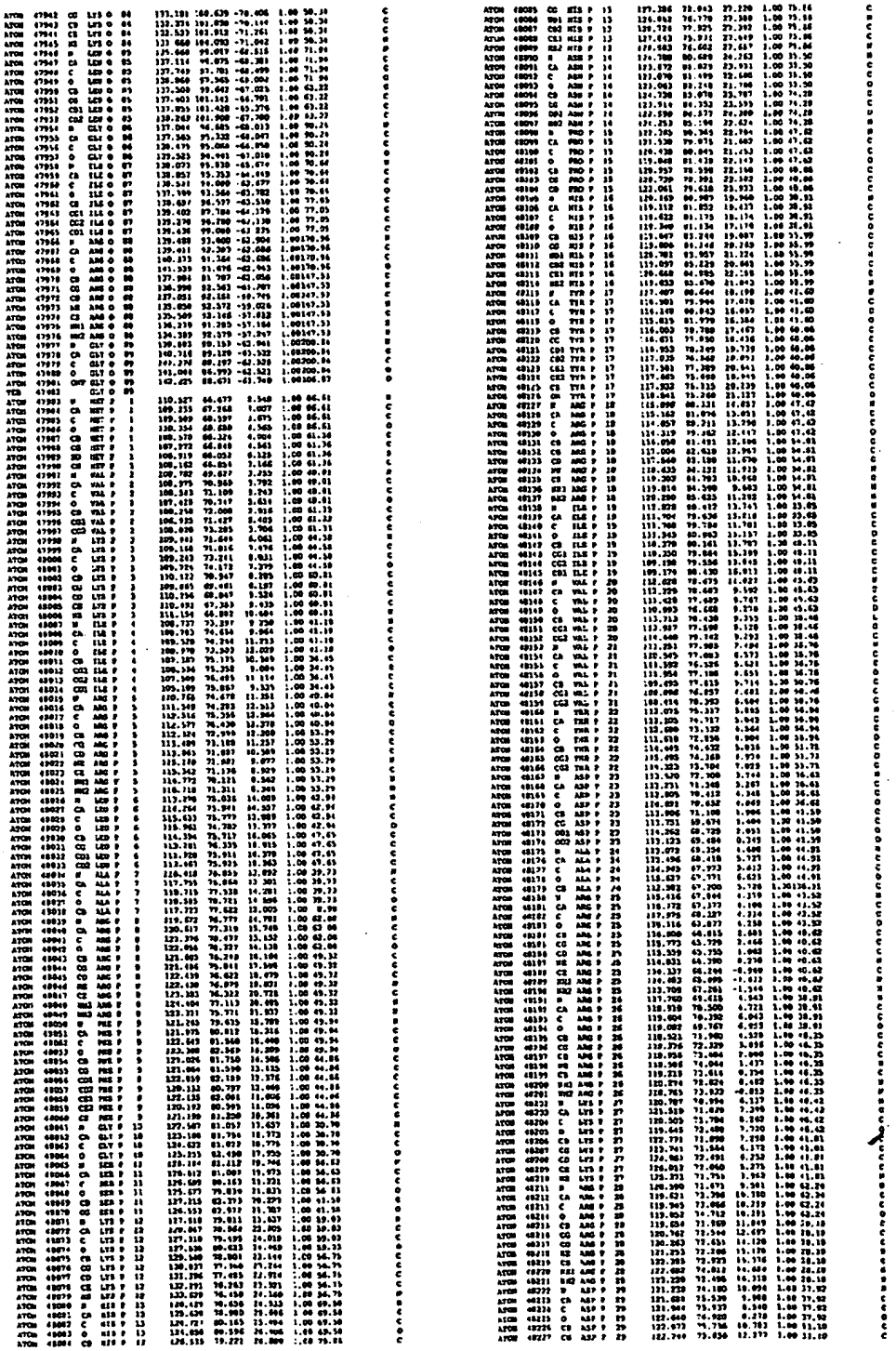


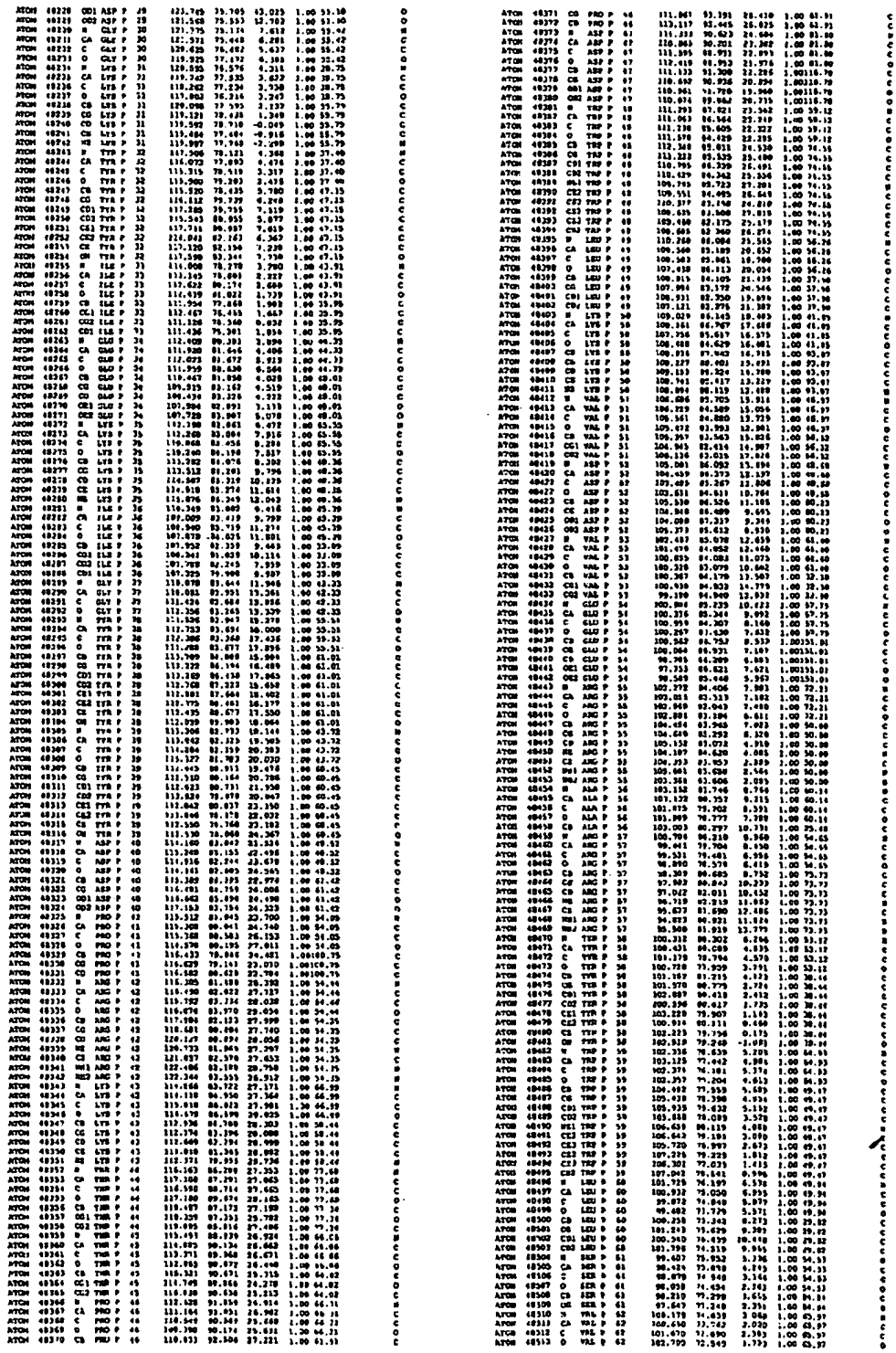


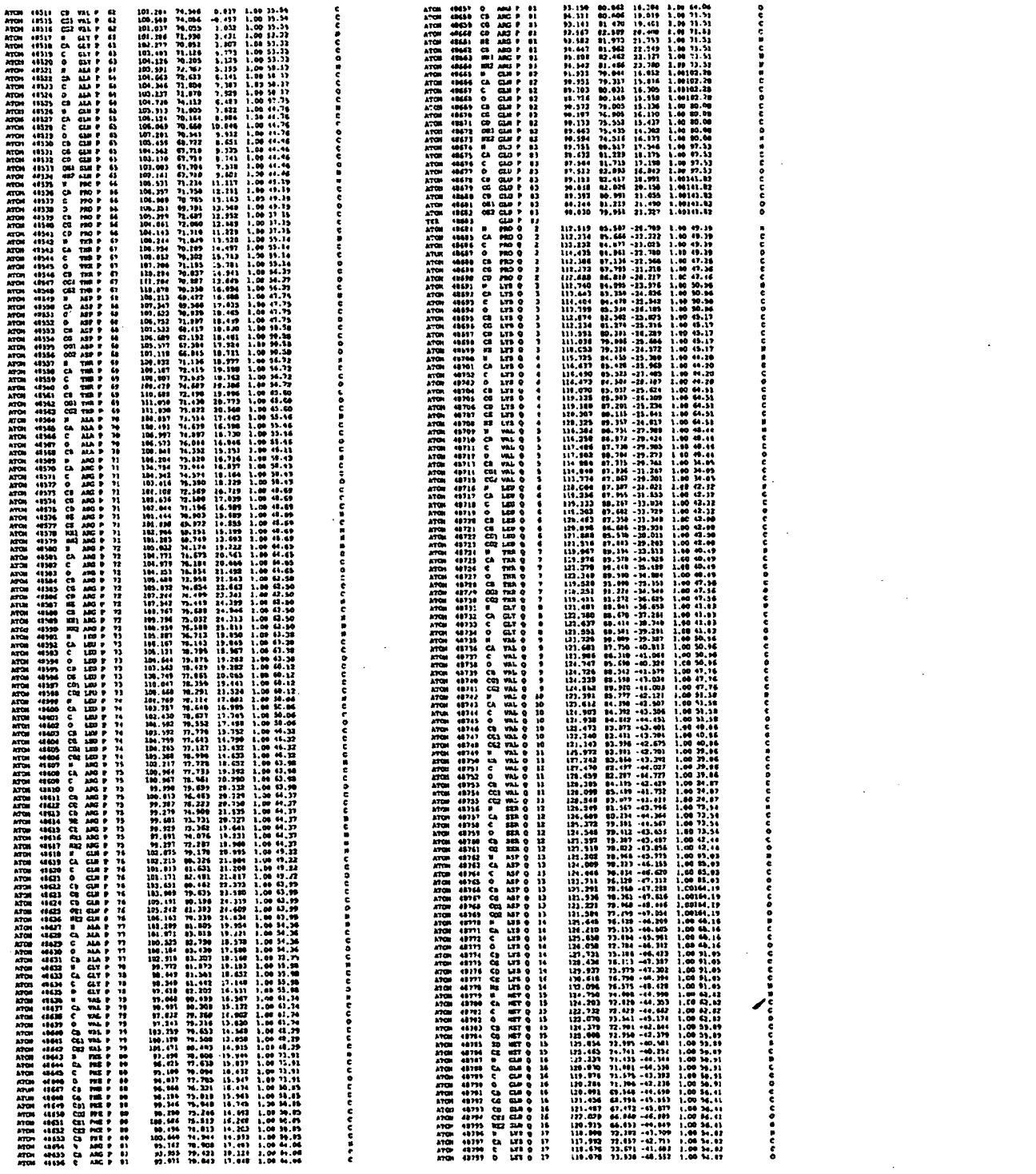


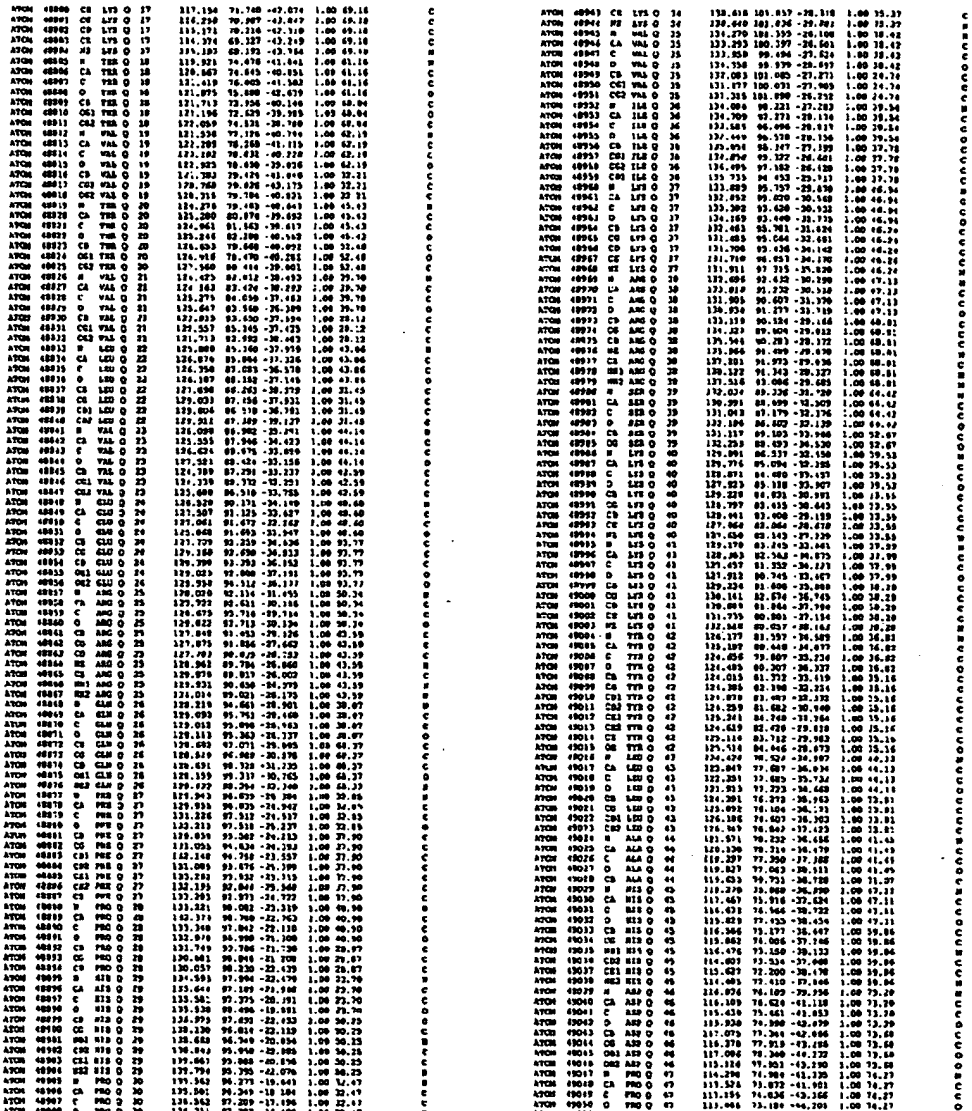


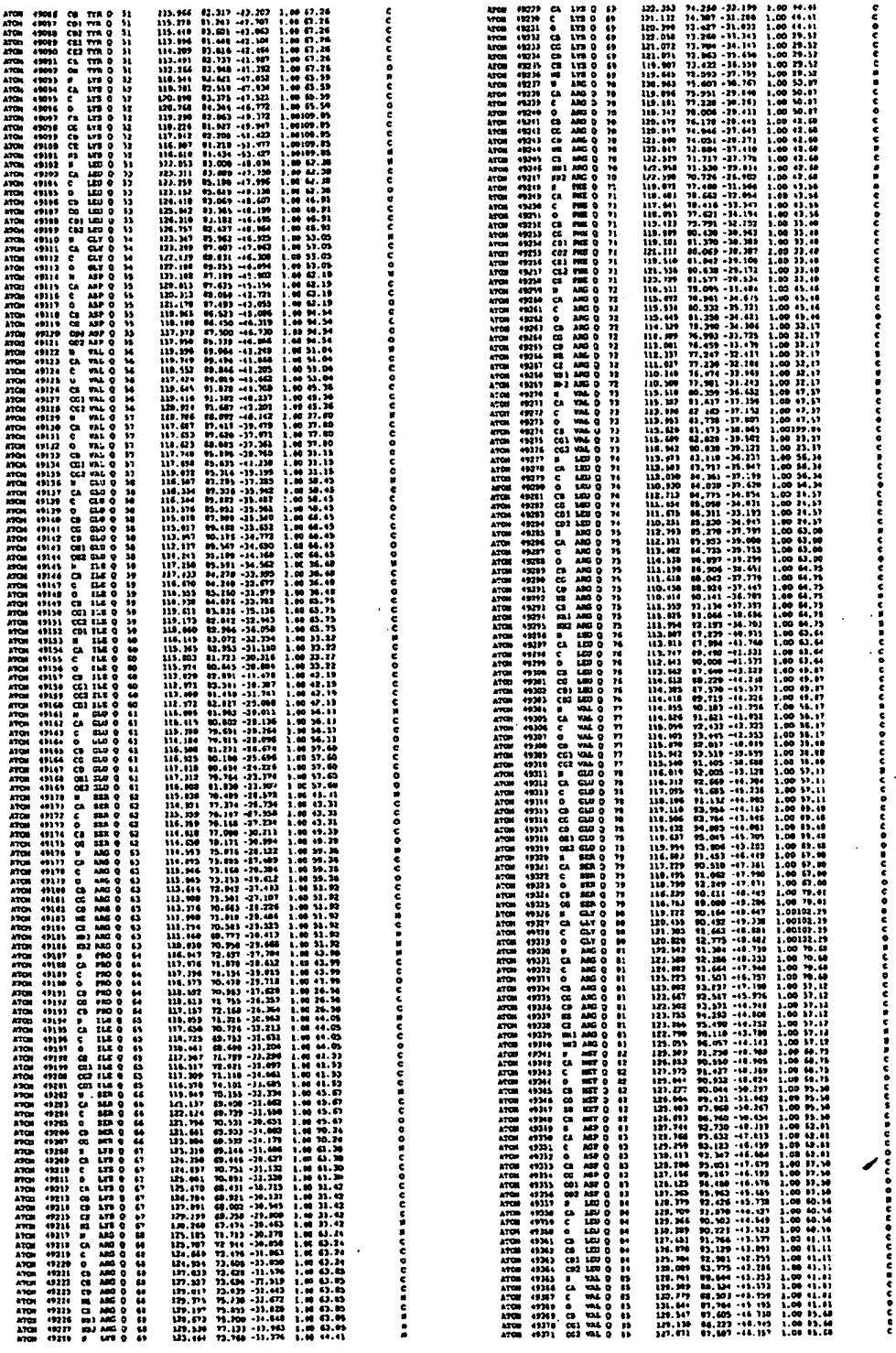


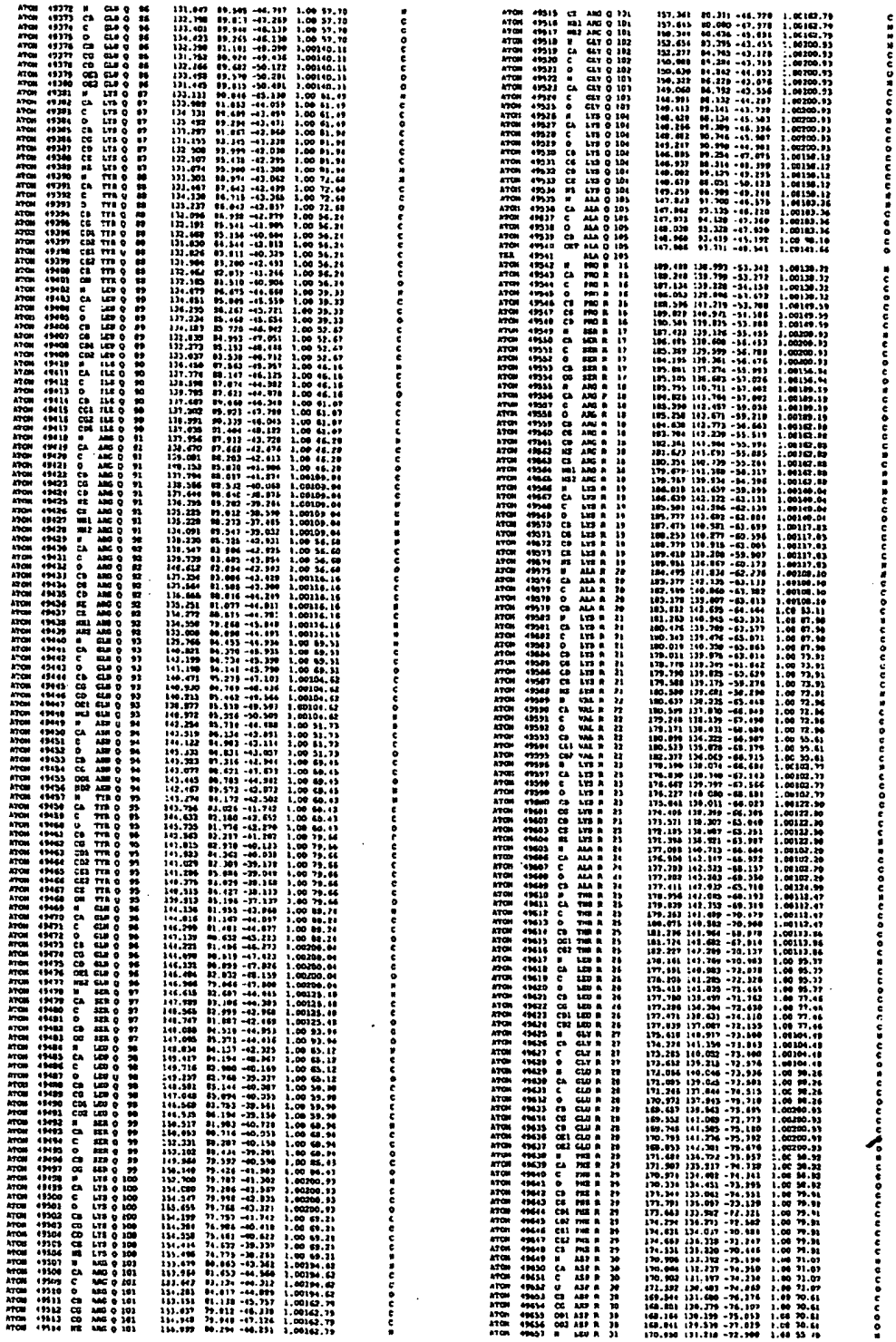


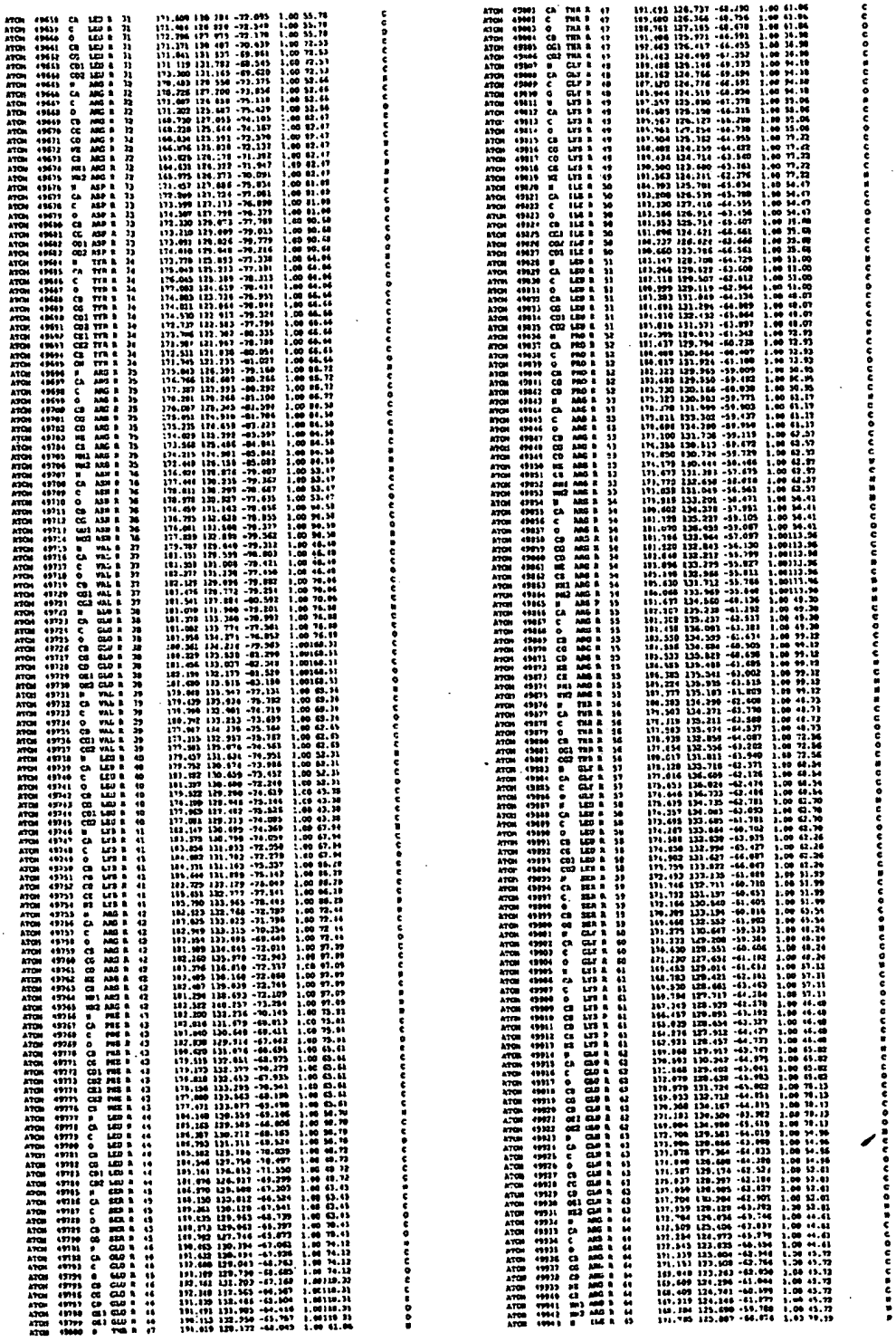


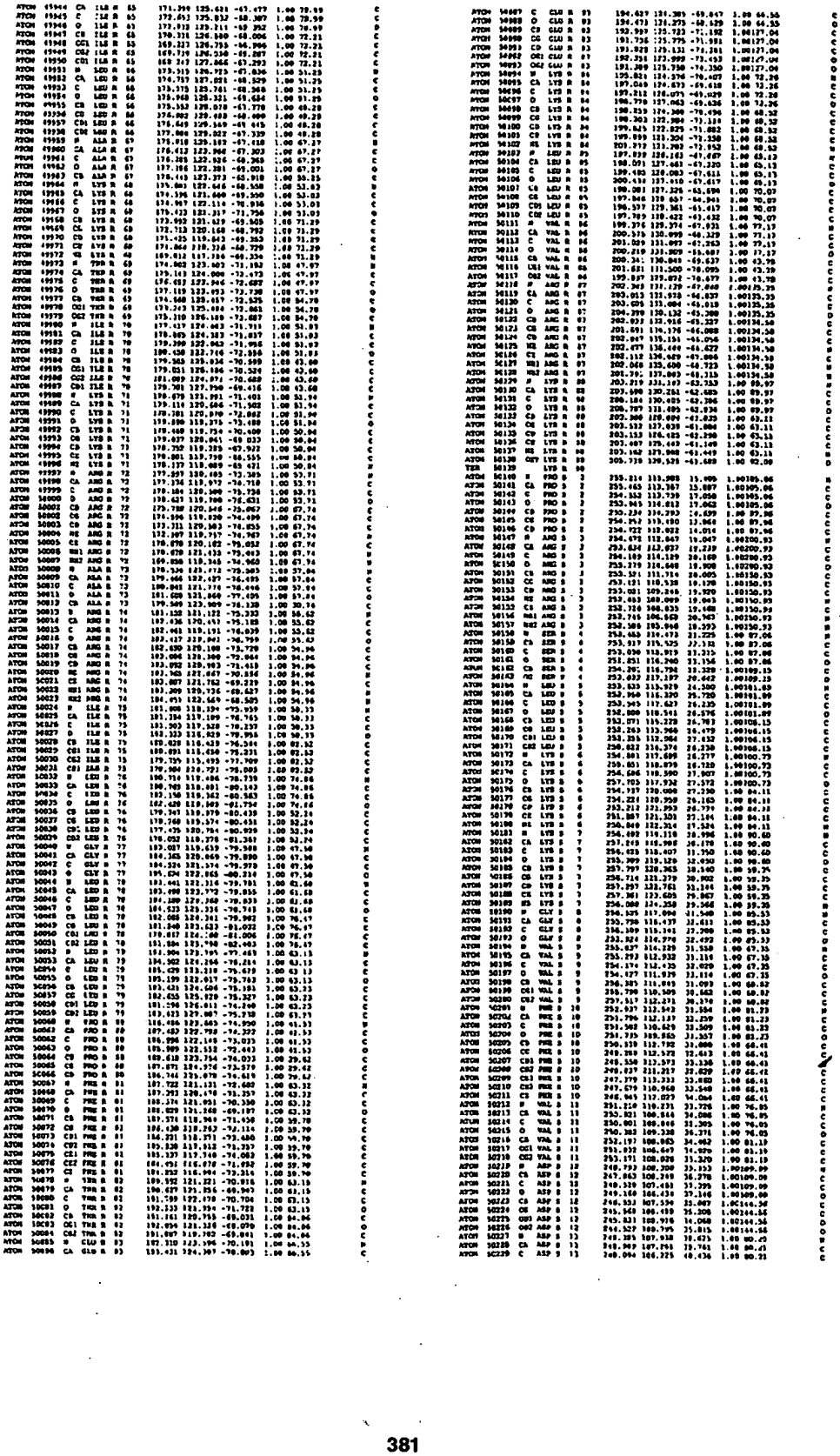


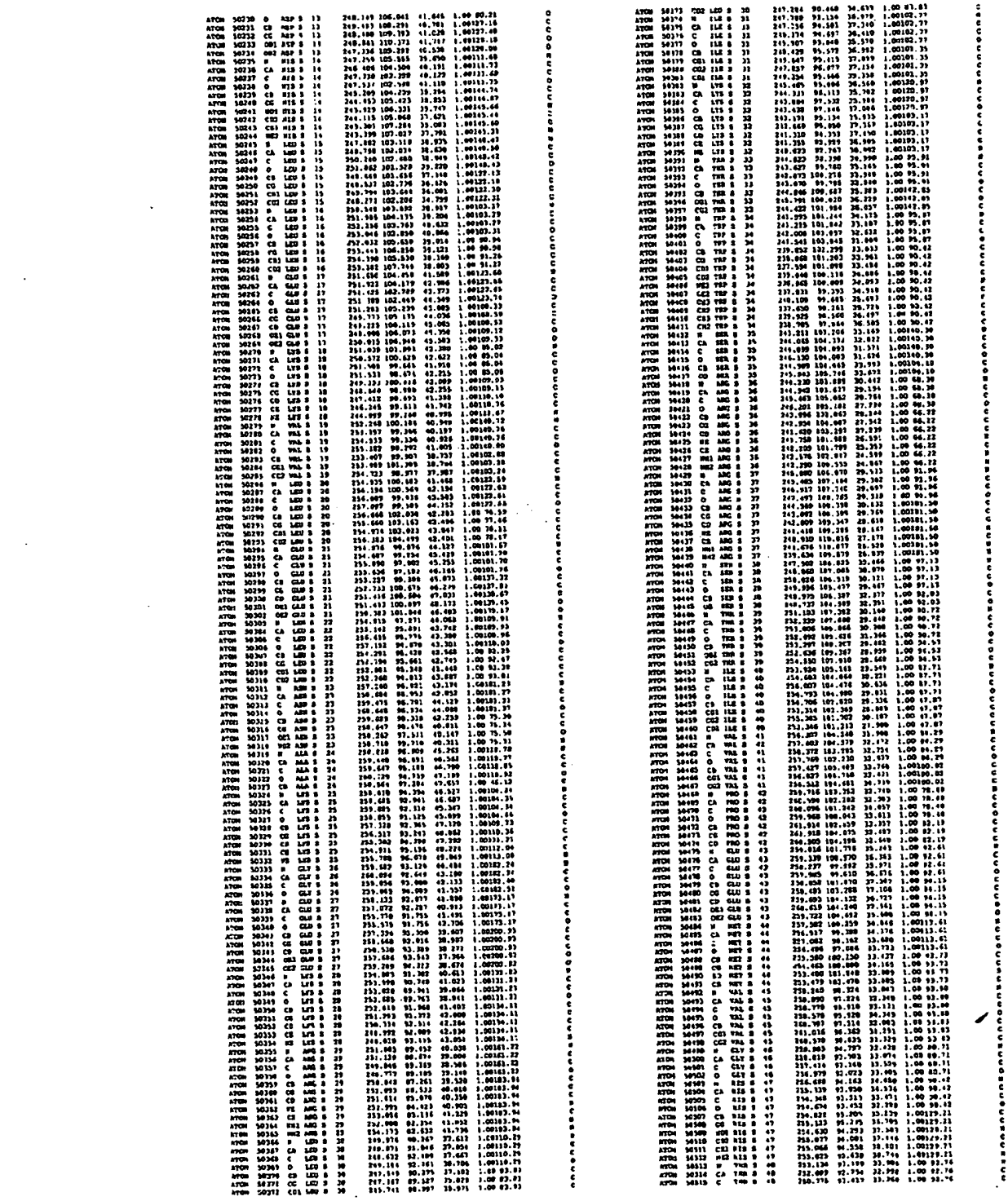


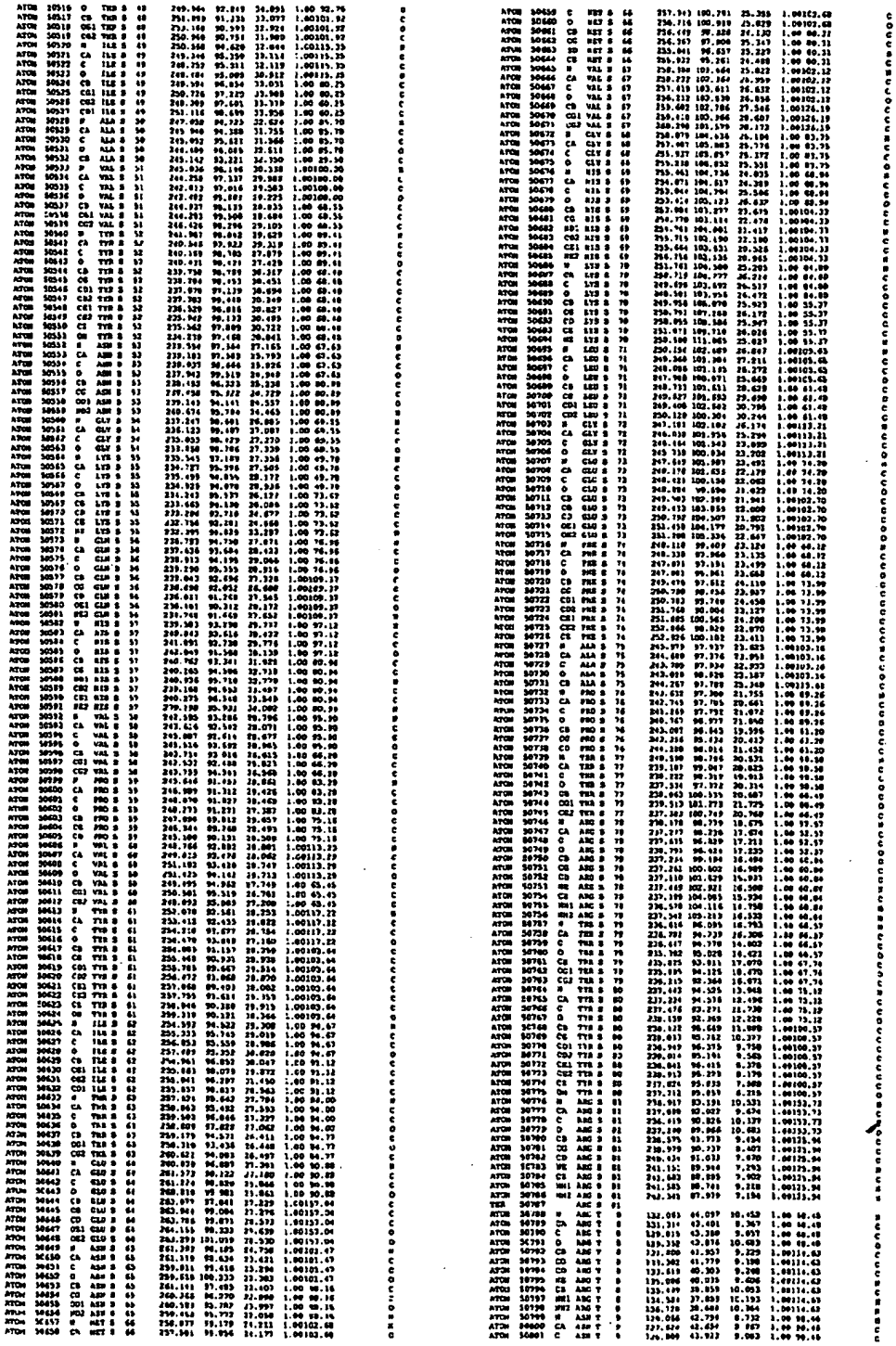


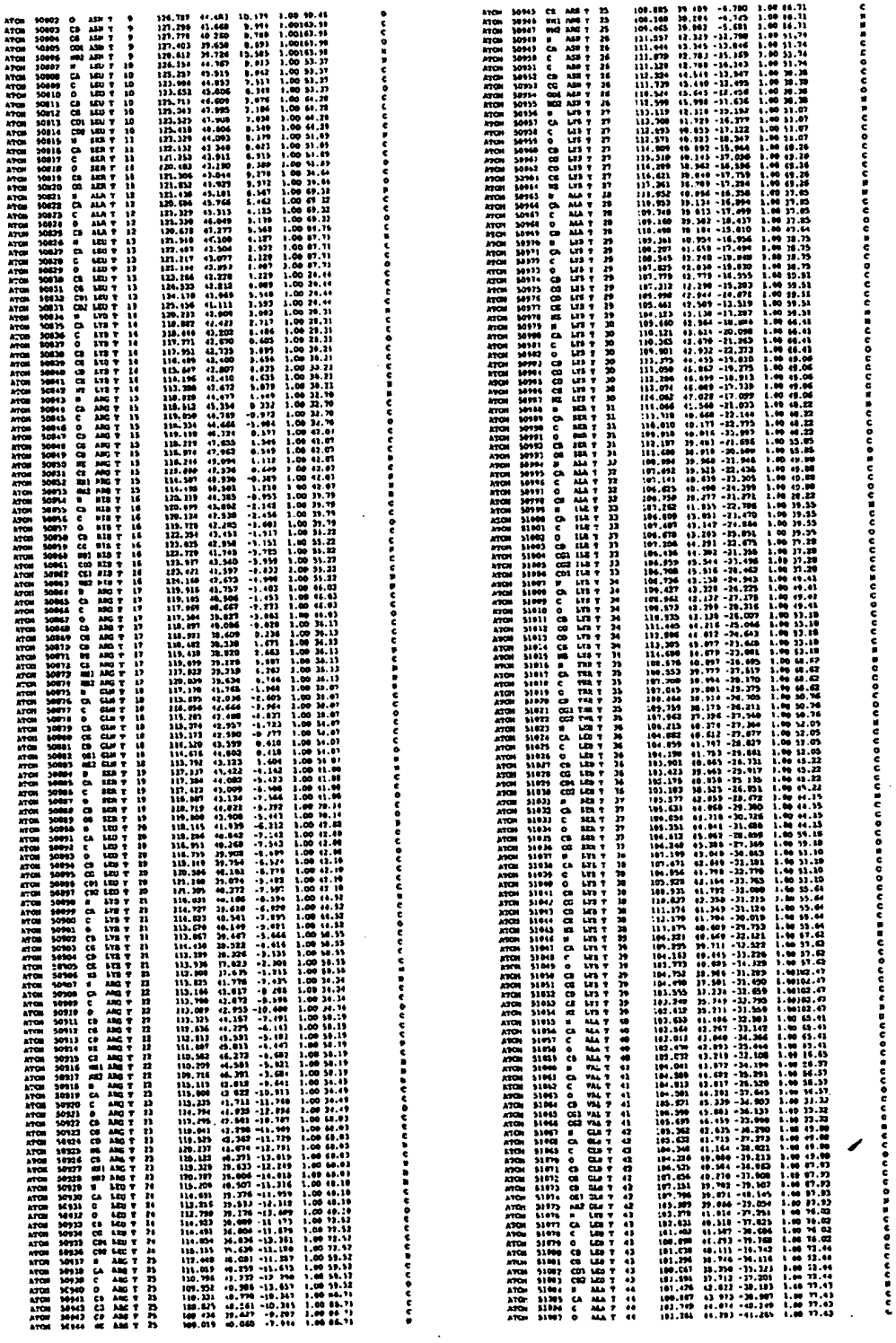


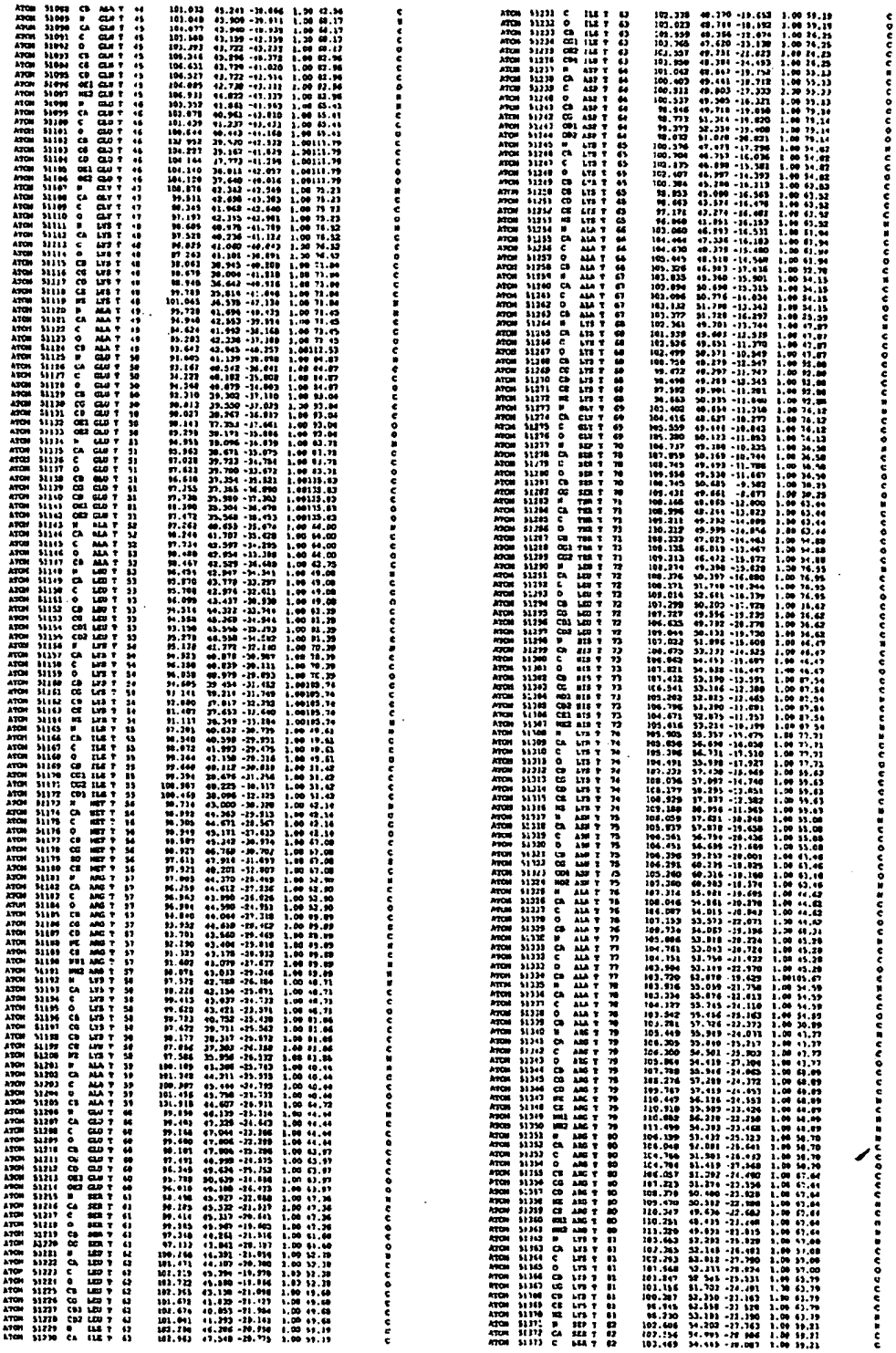


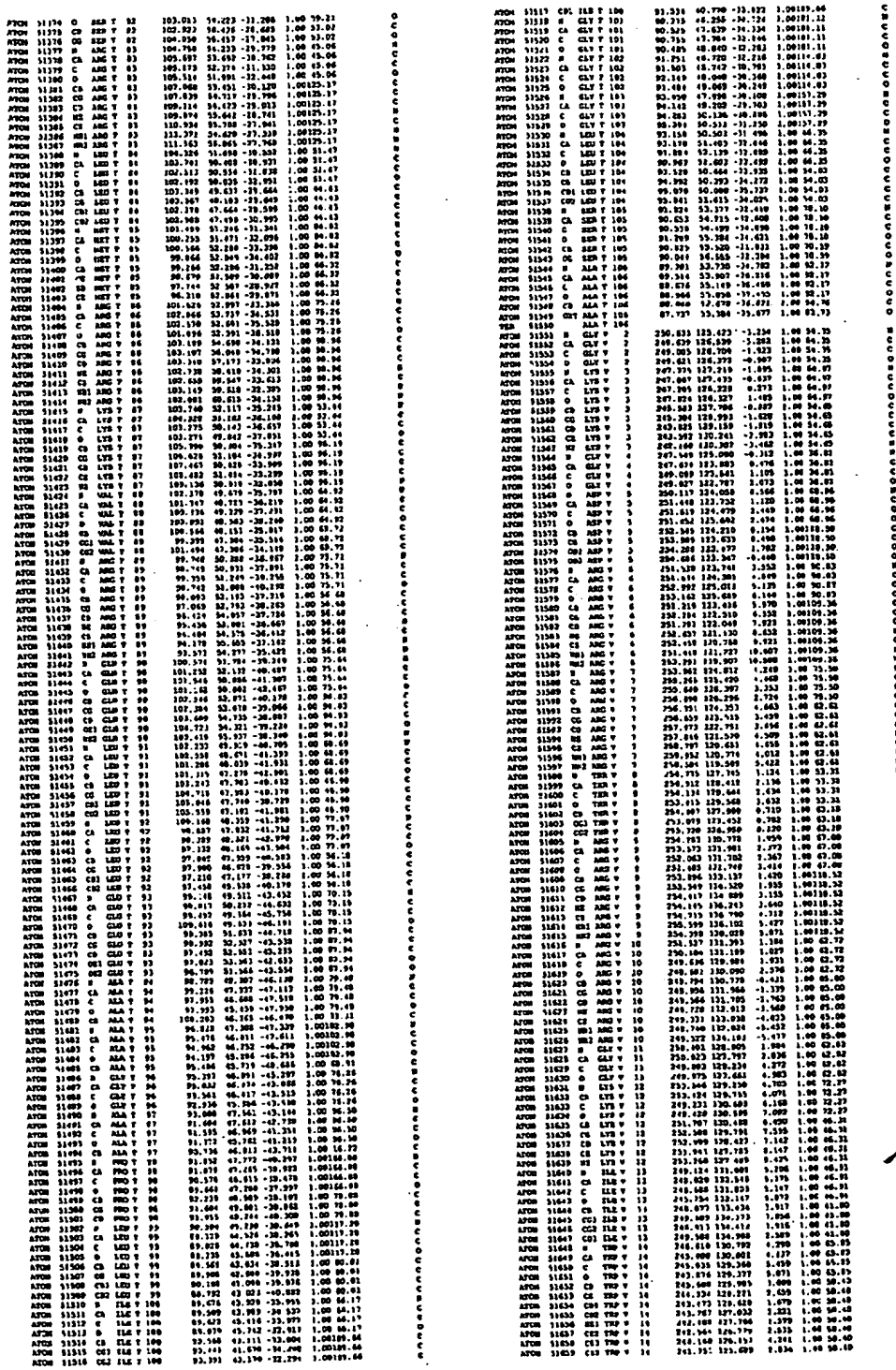


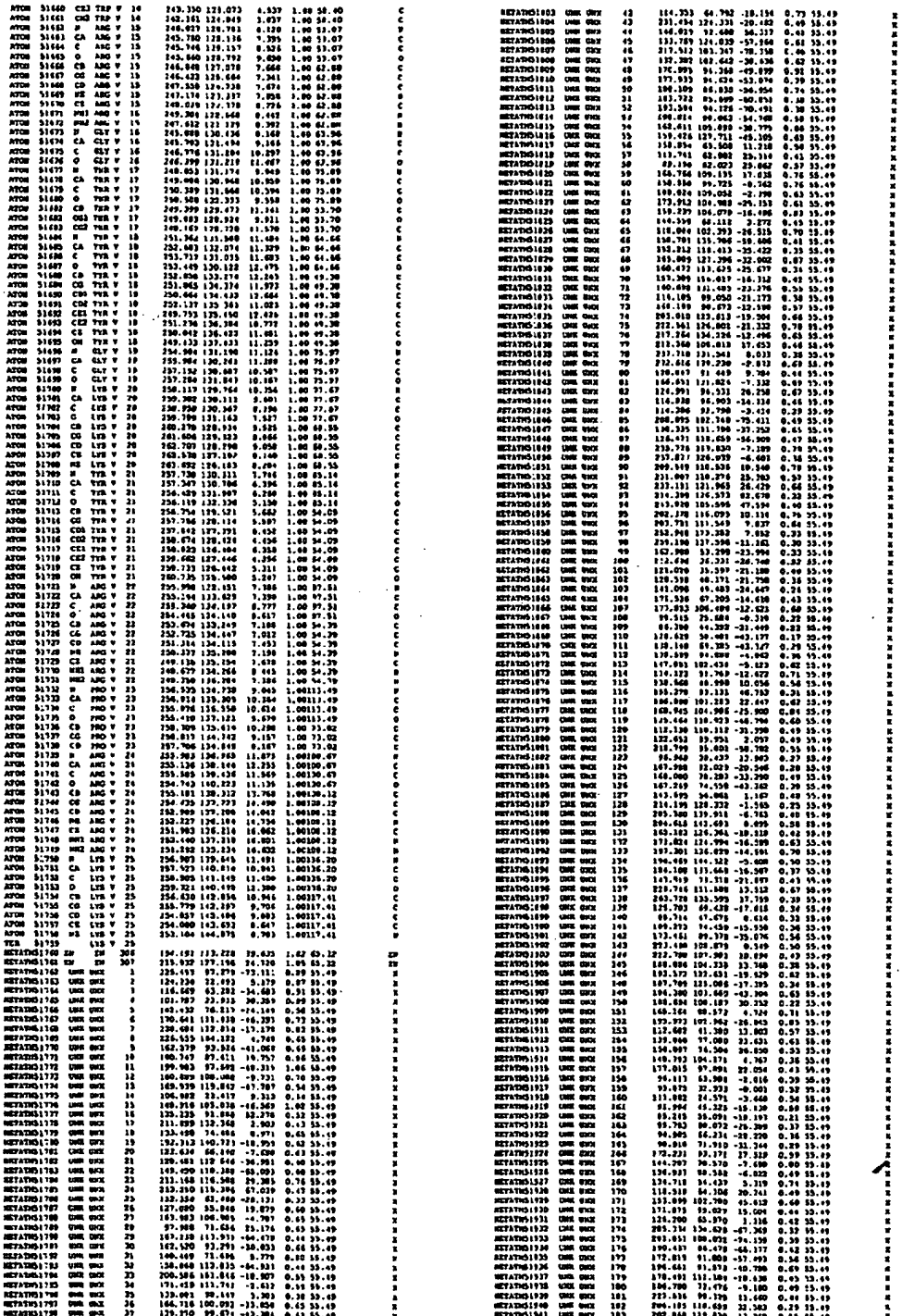












183 of 183

```
METATH51946  UNK CHX   198     195.937  95.992 -29.910  0.33 55.49          H
METATH51947  UNK UNK   200     198.601 104.580 -18.227  0.39 53.66          H
METATH51948  UNK UNK   201     201.659 108.188 -18.153  0.44 53.66          H
METATH51949  UNK CHO-  202     164.302 152.757 -31.160  0.45 55.49          H
MASTER    1379    0  190   89   89    0    0    65:1927   22    0  323
END
```

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 01 30 6060

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
Y,D	CLEMONS JR ET AL: "Structure of a bacterial 30S ribosomal subunit at 5.5 Â resolution" NATURE, vol. 400, 26 August 1999 (1999-08-26), pages 833-840, XP002182367 * See page 833 (3.6 Â), page 836 (Figure 3) and pages 838-839 (Discussion) * ---	1,2	C07K14/195 G06F17/50
Y,D	TOCILJ ET AL: "The small ribosomal subunit from Thermus thermophilus at 4.5 Â resolution: Pattern fittings and the identification of a functional site" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, USA, vol. 96, 7 December 1999 (1999-12-07), pages 14252-14257, XP002182368 * See page 14253 (Table 1) * ---	1,2	
A	AGALAROV ET AL: "Structure of the S15,S6,S18-rRNA complex: Assembly of the 30S ribosome central domain" SCIENCE, vol. 288, 7 April 2000 (2000-04-07), pages 107-112, XP002182369 * See page 108 (Figure 1 and Table 1) * ---	1-11	TECHNICAL FIELDS SEARCHED (Int.Cl.7) C07K G06F
P,X	WIMBERLY: "Structure of the 30S ribosomal subunit" NATURE, vol. 407, 21 September 2000 (2000-09-21), pages 327-339, XP002182370 * See page 332 (Figure 5) and 335 (Table 2) * -----	1-11	

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
BERLIN	13 November 2001	Korsner, S-E

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)